Concordancing Second Life discourse

Włodzimierz Sobkowiak (SL=Wlodek Barbosa)

Abstract

Second Life is currently the most popular MUVE (Multi-User Virtual Environment) in existence, with over 10 million registered residents, most of them non-native speakers of English, which meanwhile remains SL's lingua franca. SL discourse is in many ways similar to that of Real Life (RL), but there are important differences, especially on the lexical, phraseological and stylistic level which are due to the need to communicate in-world about objects, concepts, events, activities and skills specific to this virtual world. New SL residents who are EFL (English as a Foreign Language) learners at the same time could benefit from resources specifically targetting their needs, both SL- and EFL-related. So far no such tools and materials exist which could be used for EFL teaching and learning in SL. This text proposes to use existing corpus analysis software to concordance SL discourse and extract SL-related vocabulary and phrases which could then be used for EFL pedagogy in SL. An example of such a procedure using *The complete fool's guide to Second Life* by Foolish Frost (SL identity) as analysed text is presented.

1. Introduction

With over 10 million residents, *Second Life* has now become the most popular MUVE in existence. MUVE refers to online, multi-user virtual environments, sometimes called virtual worlds or (incorrectly) MMORPGs — Massively-Multiplayer Online Role Playing Games. "Second Life (SL) is an internet-based virtual world launched in 2003, developed by Linden Research, Inc (commonly referred to as Linden Lab), which came to international attention via mainstream news media in late 2006 and early 2007" (Wikipedia). Residents, appearing in the form of animated avatars, can communicate by standard text-chat and (since July 2007) by voice in order to make friends/love, entertain, do business or follow a variety of courses on offer.

Focusing on the latter, "there are now many universities, colleges, schools and other educational institutions researching the use of Second Life as an environment for teaching and learning which offers a community of practice and situated constructivist learning" (Wikipedia). *Second Life* is now arguably the most technologically advanced platform to use as a Learning Management System (LMS) to afford interactive, immersive, communicative foreign language learning and teaching. Considering that English functions as an unchallenged lingua franca in SL, the potential of SL for English as a Foreign Language (EFL) teaching and learning is hard to overestimate. A number of groups/organizations/companies/schools have already sprung up to life in SL which offer EFL tuition in a variety of forms and settings: *SLEnglish* (http://www.secondlifeenglish.com/news.php) situated on the *English Village* (http://www.eslteacherlink.com/) island and *LanguageLab* (http://www.languagelab.com/index/) occupying a number of SL islands are just two out of many examples.

An important task of these schools is to provide SL 'newbies' (new residents) who are not native speakers of English with EFL tuition specifically geared towards their newbie needs,

such as locomotion, avatar appearance, shopping, searching, map reading and socializing in the virtual environment. And, according to the recent statistics (September 9th 2007), "60% of the SL residents don't speak English as their first language" (http://gwynethllewelyn.net/2007/09/09/the-multi-cultural-second-life%C2%AE-and-penetration-rates/). "Interestingly enough, the language spread in SL is quite different from what happens in the real world! According to the Wikipedia, the first ten languages in the world are, respectively, Mandarin Chinese, Spanish, English, Arabic, Hindi, Portuguese, Bengali, Russian, Japanese and German. Compare that to Second Life's English, Portuguese, German, Japanese, French, Italian, Dutch, Spanish, Polish and Mandarin Chinese!" (ibidem; emphasis – WS).

While there are many textual resources in English, both online (electronic) and off (printed traditionally) catering for such needs, I believe they can be exploited in more innovative, and also more linguistically revealing, ways than has so far been the case. The following is an attempt to demonstrate how one such SL introductory guide text written for SL newbies can be processed through a rather standard suite of concordancing scripts to reveal immediately teachable SL-related English vocabulary and phraseology items.

The text is *The complete fool's guide to Second Life* (henceforth: "the Guide"), by Foolish Frost (SL identity), to be found here: http://www.sldrama.com/index.php?page=2 (used with permission). Because the text is stylistically a mixture of two very different discourses, namely an ordinary guide and a narrative about the Author's newbie experience (see Appendix for samples of both), I decided to delete all of the latter text (conveniently placed in insets of the original pdf document). The resulting text, i.e. the guide proper, now counts 7468 words, 58 paragraphs and 32639 characters. The number of word-types stands at 1359, and the token/type ratio = 5.5.

2. Concordancing *The complete fool's guide*

An evaluation version of PhraseContext by Hans J. Klarskov Mortensen (http://www.hjkm.dk/PhraseContext/), ver. 1.0.2., was used to process the text. This is a relatively simple to use, yet powerful, package with standard concordancing functions, such as KWIC, wordlist generation, word-cluster (within 3-6 word range) and collocate extraction, t-score, z-score and MI statistics. These functionalities are quite adequate for the average needs of an EFL teacher.

2.1. Words

Expectedly, the frequency-ordered wordlist generated from the Guide begins with the usual bunch of function words: *the, to, you, and, a, of, in, is, can, that, your, are, will, it, on*. The first intrusion comes at rank-position 16 – *SL*, i.e. the acronym of *Second Life*, with 60 occurrences. But *Second Life* is of course notorious for some of its key vocabulary, such as *avatar, teleport, prim, griefer, rez, sim, sandbox*, which reflect the many unique concepts which newbies must acquire before they can comfortably function in the virtual environment. To easily extract these words from the rest of 'ordinary' vocabulary I confronted the Guide wordlist with that of BNC (the so-called 'context-governed spoken material'; 6.2M tokens, 79,906 types; see http://www.kilgarriff.co.uk/bnc-readme.html). The result, i.e. words absent from the BNC wordlist, are listed below in Figure 1, with their frequencies in the Guide. Evident noise and errors have been removed. For example, the hapaxes *get's* and *layed* in the Guide wordlist were due to these sentences in the Guide, respectively: *Anything that get's*

returned to you from a person's land will return to here and *The folders are layed out in a standard branching format...*

Figure 1. Words unique to the Guide (as against BNC), with their frequencies

avatar 39, teleport 16, avatars 12, prims 11, lindens 5, animations 4, avatar's 4, chatbar 4, griefer 4, mouselook 4, notecards 4, trash 4, 3d 3, griefers 3, rez 3, ctrl 2, http 2, mentors 2, rebooting 2, rezzed 2, secondlife 2, sims 2, uploaded 2, website 2, alot 1, animates 1, browser 1, cached 1, checkbox 1, checkboxes 1, contests 1, detects 1, emote 1, framerate 1, gray 1, grays 1, griefing 1, internet 1, jewelry 1, login 1, Mahulu 1, maneuvering 1, neighbors 1, notecard 1, offline 1, php 1, premeditatedly 1, RL 1, sandbox 1, scripting 1, scrollbar 1, scrolling 1, slider 1, slowdowns 1, submenu 1, teleported 1, teleporting 1.

The reader will appreciate that this is no place to explain SL to the uninitiated, so – for the sake of this discussion – I will take just the first two words from the list, which happen to be the two quite central to the very idea of SL and ones of immediate need to a complete newbie: *avatar* and *teleport* (interestingly, the word *newbie* does not appear even once in the Guide; this may have to do with its slightly condescending tone). The former denotes the SL 'self' of the resident, an animated personalized figure with a unique identity; the latter – one of the methods of locomotion available in SL, on top of walking, running and flying. SL residents tend to manipulate their avatars a great deal in a variety of ways, and they teleport frequently.

A few further clarifying comments are in order. *Lindens* refers either to the (employees of the) *Linden Lab* or the SL currency, linden dollars. Some words in Figure 1 denote various elements of the SL Graphical User Interface, such as *chatbar* or *scrollbar*, or general MSWindows functionalities, such as *ctrl* (the control key). Some other words are missing from BNC quite simply because at the time when it was collected and when its canonical form was set, i.e. between 1991 and 1994 (http://www.natcorp.ox.ac.uk/corpus/), they did not yet function in public discourse, e.g. *Internet* or *website*. Some Guide words are missing from BNC because they appear in the former in their American spelling. Finally, *Mahulu* is a Real-Life (RL) place in the Republic of Congo, which is simulated in SL.

In Figure 2. a complete KWIC concordance of the word *avatar* is presented, sorted on the last word of the left context, thus exposing its left-hand-side collocates.

Figure 2. KWIC concordance for "avatar", sorted on the left collocate

```
eating one yourself or buying an avatar design off of one of your fellow
rt! You will sometimes notice an avatar with the last name of Linden. Th
s used to remove the highlighted Avatar from the list. Offer Teleport se
rous option. It allows the other avatar you chose to edit and copy ALL o
arts that create a certain style avatar for easy access. That would incl
n get more information about the Avatar and the person behind it. Remove
ing too. You could even have the avatar automatically say "Ho, ho, ho!"
s, the clothing is a part of the avatar, in most cases you store the clo
r Teleport sends an offer to the Avatar that allows them to Teleport dir
when you first login. Unlike the avatar of the story, you may see other
 can respond to you.   While the Avatar teleported to the mainland of Se
now how to chat in SL yet.   The Avatar just received his first Notecard
h for you to drop items on their avatar. It's a good idea to ask a perso
 the mouse to look around. This 'avatar eye view' has several advantages
ed Sim can be slow to respond to avatar actions, such as chat or movemen
make a gesture that animates you avatar to look like he's laughing and a
```

```
 restore your camera is move you avatar with the arrow keys. If you are
mple. To change the shape of you avatar, you use the slider controls. Th
to create pre-build actions your avatar can do. You could make a gesture
shut down the SL client and your avatar will fade from the world of Seco
cts and Gestures to animate your avatar with poses. Dancing, bowing, sta
 avatars visual appearance. Your avatar can look like anything, and ofte
a small map that rotates as your avatar turns his facing left and right.
extures to color and detail your avatar. These images can be applied to
 want to say and hit enter. Your avatar will then say the chat to everyo
n until you hit the ground. Your avatar will exit flight mode and begin
le, while the yellow dot is your avatar.   Let me point out something th
 play back in order to make your avatar 'emote' to others. You can even
have to stand and then move your avatar. Either way, your camera will sn
le most important aspect of your avatar, second only to the ability to c
ter.   Changing the look of your avatar can seem intimidating at first.
for the current position of your avatar. This landmark will be stored in
ance! First, right click on your avatar and choose Appearance. You can a
t here to change clothes on your avatar as your mood changes. Gestures:
eft-clicking and holding on your avatar (your character in-world) in the
lry, and even body parts on your avatar. To begin building, you can go t
no further info to register your avatar. Basic accounts do not have the
you. This allows you to see your avatar and your surroundings in an easy
sual representation of what your avatar will look like if you slide the
```

It will be seen that by far the most frequent left collocate of *avatar* is *your*[1]. This is an expected outcome in a text which is explicitly a guide for the newbie SL resident. From the pedagogical perspective it is worth pointing out to the naturally occurring example of linking /r/, a connected-speech phenomenon of some difficulty to EFL beginners and intermediates, especially those aiming at standard British pronunciation (the so-called RP). Incidentally, there are quite a few linking /r/'s originating on the other end of *avatar* in those contexts where the following word begins with a vowel. It is easy enough to order the concordance on the right context, of course, to expose these lines. This is not done here for reasons of space.

The predominance of the definite article over the indefinite before *avatar* can also be exploited didactically, as well as the clearly adjectival use of the noun in *avatar eye view*.

To take the other key term, a KWIC concordance has been made for *teleport*, with its inflections, this time sorted on the right collocate. The results appear in Figure 3.

Figure 3. KWIC concordance for "teleport", sorted on the right collocate

```
 a landmark, you can just click 'Teleport' and you're on your way. You w
o the Avatar that allows them to Teleport directly to you. Pay send Lind
 choosing the options to Freeze, Teleport home, or even Ban the griefer.
l you have to do is agree to the teleport offer. In that case, you will
o sent the invitation. Sending a Teleport offer is made easy by SL. You
ndmark folder in your inventory. Teleport offers from other people are e
 lot of frustration.   Accepting teleport offers is fairly simple: If yo
is made easy by SL. You can send Teleport offers to others from their pr
herever you have set your 'home' teleport point to. This means that no m
hted Avatar from the list. Offer Teleport sends an offer to the Avatar t
around you as you move about and teleport. The fact is, SL is a dynamic
rk is an item that allows you to teleport to a place, much like the 'tic
```

---

[1] Three errors by Foolish Frost become immediately visible in this listing: "animates you avatar", "move you avatar" and "shape of you avatar".

```
   of people in one place when you teleport to a Sim with an Event. That's
   format. Using this map, you can teleport to new locations just by doubl
apparent reason. You will try to Teleport to somewhere only to have it f
spond to you.    While the Avatar teleported to the mainland of Second Li
s a good idea to go there before Teleporting to the mainland. Once in th
   an event, all you have to do is teleport using the option at the bottom
```

It will be noticed, that – unlike *avatar* – *teleport* functions in two morphosyntactic capacities: as verb and noun. In the former, there are examples of rather common usage with the leading infinitival 'to' particle or a modal: *you can teleport to new locations*. For more advanced EFL students the *to*-less use of the verb in the last line of the concordance may stimulate discussion on the syntax of infinitives, emphasis, extraposition, etc., for example: if "clauses which explain the exact meaning of *do* <u>can</u> have the infinitive without *to*" (Swan 1980:321; my emphasis – WS), how often, on average, <u>do</u> they, and are there any rules about when one can drop the *to*?

For *teleport* in its nominal function, the compound *teleport offer* appears to be a common collocation. This is due to the pragmatics of communication and locomotion in SL: avatars often offer teleports to their friends to (instantaneously) bring them to the place they are currently in. Thus, both teleport offers and requests are common, the relatively lower incidence of requests probably being ascribable to the universal pragmatics of interpersonal communication (e.g. fear of imposition). Thus, avatars *send, accept* and *agree to* teleport offers – all these collocations are exemplified in the concordance and can be practised by the EFL learning SL newbies.

At his point it is rather obvious that concordancing SL discourse can reveal much more than simply frequent SL-specific terminology. It is possible to scan for typical collocations of given keywords through long concordance lists, but it is of course more expedient to use software dedicated to extracting frequent phrases from text directly. *PhraseContext*, as well as most text-analysis and concordancing software nowadays, has such functionalities. In what follows I will demonstrate how they can be used to find EFL-useful common multi-word phrases in SL discourse.

2.2. Phrases

PhraseContext will extract phrases of between two and six words, and arrange them by frequency of occurrence in the input text. The EFL teacher in Second Life thus has immediate access to what many learners find the hardest nut to crack in English: its incredibly rich phraseology, with phrasal verbs, complex noun phrases, multiple compounds, and the like. As will be seen below, some of the commonest phrases in the Guide turn out to have little SL specificity, but they do remain extremely useful to learners for their general command of conversational English.

In a text of the size of the Guide one would not expect a significant number of 6-word phrases. There are a few, however, of shorter length. These are collected in Figure 4, together with their respective frequencies. It is obvious that some common shorter phrases are inevitably contained in longer ones, for example *a good idea to* in *it's a good idea to*. These were omitted in the listing.

Figure 4. Word clusters in the Guide, sorted by frequency

5 words		4 words		3 words	
8	*it's a good idea to*	5	*allows you to edit*	11	*you want to*
3	*all you have to do*	5	*the bottom of the*	6	*a lot of*
		4	*allows you to see*	6	*all of the*
		4	*this allows you to*	6	*one of the*
		3	*at the bottom of*	6	*you can also*
		3	*have the ability to*	6	*you can see*
		3	*that allows you to*	5	*up and down*
		3	*the terms of service*		
		3	*the top of the*		

A number of interesting observations can be made of these clusters. First, they are all <u>phrases</u> in the syntactic sense of the word, rather than accidental tuples of words, as could be the case considering that the software's intelligence does not extend beyond mere statistics. Second, these phrases are all potentially useful in the context of EFL teaching/learning, both in SL and in RL. Third, some of them are of course rather trivial grammatically, and thus not much fodder for fruitful pedagogical exploitation, but (fourth) there are ones which tend to be syntactically and/or pragmatically troublesome to learners. Take *all you have to do*, for example, which appeared before in the concordancing context as an example of an extraposed nominal clause under emphasis (or, strictly speaking, thematization/topicalization, see http://en.wikipedia.org/wiki/Topic-comment) triggering the loss of the infinitival particle *to*. Quite apart from the *to*-less infinitive, there is also the difficult semantic contrast of *all* versus *everything* and other such indefinite pronouns. A simple Google search shows that *all you have to do* is by far more frequent on the net (2,880,000) than *everything you have to do* (32,800), but the relatively numerous incidence of the latter shows that it is not simply an error. The (inquisitive) learner will now ask about the rules governing this variation, if any.

The *allows you to* or *you want to* phrases are visibly very common for reasons which have to do with the pragmatics of instructional text. As such they are of immediate use to the newbie in the SL context, as well as to the EFL learner outside it. While the *you* in the Guide is <u>not</u> the generic type, because Foolish Frost addresses each and every newbie directly in his text, the issue of general reference with *you* as well as *one* (even *man* for the advanced and interested) could be raised by the teacher in this connection.

While *have the ability to* may be syntactically unproblematic, it is still exploitable pedagogically on the pragmatic plane, for example in its subtle contrast to *be able to* (about a hundred times more frequent on the web). Finally, *at the bottom of*, *at the top of* and *up and down* are common enough adverbials, again of use not only in the SL virtual world, offering ample opportunities for contextualized communicative language practice.

The phrases in Figure 4. afford a variety of pronunciation exercises as well, especially in the so-called sandhi, i.e. the phonetics of inter-word juncture. To name just a few points: the voicing of the final /v/ in *all you have to do* and *the terms of service*, vocalic hiatus in *allows you to edit*, /t/ dentalization in *at the bottom of*, plosive gemination in *you want to*, schwa deletion + nasal syllabification in *you can also* or alveolar stop deletion in *up and down*.

3. Conclusions

EFL pedagogy in Second Life is very young. It is enough to scan the archives of the relevant discussion lists, wikis and blogs, such as https://lists.secondlife.com/cgi-bin/mailman/listinfo/educators or http://www.secondlifeenglish.com/news.php to witness educators trying to fathom the new pedagogical possibilities, problems and dangers generated by this thoroughly interactive, immersive, communicative and social virtual environment. There is little solid research at the interface of virtual worlds and education (e.g. Coffman & Klinger 2008), let alone Second Life and EFL. It seems that, while text concordancing and corpus analysis have by now found a secure place in Real Life EFL teaching and learning (e.g. Aston 2001, Leńko-Szymańska 2005, Scott & Tribble 2006), Second Life is still pristine in this respect. It is of course possible to use standard corpus/concordance facilities, resources and materials within SL as well as outside it. Considering the lexical, phraseological and stylistic novelty of the environment and its associated discourse, however, it is only fitting that it should be explored in all its richness for pedagogical purposes. More and more EFL learners will be going into Second Life to acquire or improve their skills, both linguistic and SL-related. Concordancing SL discourse will provide tools and resources to help them do this.

References

Aston,G. (ed.) 2001. *Learning with corpora*. Houston, TX: Athelstan.

Coffman,T. & M.B.Klinger. 2008. "Utilizing virtual worlds in education: the implications for practice". *International Journal of Social Sciences* 2.1.29-33.

Foolish Frost. (n.d.) *The complete fool's guide to Second Life*. http://www.sldrama.com/index.php?page=2

Kilgarriff,A. 1996. *BNC database and word frequency lists*. http://www.kilgarriff.co.uk/bnc-readme.html.

Leńko-Szymańska,A. 2005. "Korpusy w nauczaniu języków obcych". In B.Lewandowska-Tomaszczyk (ed.). 2005. *Podstawy językoznawstwa korpusowego*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego. 221-39.

Scott,M. & C.Tribble. 2006. *Textual patterns. Key words and corpus analysis in language education*. Amsterdam: Benjamins.

Swan,M. 1980. *Practical English usage*. Oxford: Oxford University Press.

Appendix. Samples of a SL-newbie narrative and guide from *The complete fool's guide to Second Life* by Foolish Frost

Narrative:

I must have blacked out, and was looking around for several moments before I figured out I wasn't sitting at my desk anymore. I wasn't even in my house anymore... Dear Lord... I wasn't in MY BODY anymore... I didn't know whose body I had gotten, but it was wired weird compared to mine. My hands weren't mine anymore, and my body felt funny in ways I could not even begin to describe. Even stranger, while standing didn't take much effort, I couldn't

figure out how to sit down. Considering the shock, I did what any person would do under the circumstances. I panicked.

Guide:

First, go to http://secondlife.com/ and click on the "Join Now" button. From here, it will ask you to fill out a registration form. This form allows you to enter a first name for yourself, and to choose a second name from a list. The list normally contains dozens of last names that are currently available, but these do change over time. Some names are retired, and new ones are added to replace them as enough people register to take up all the slots available for that last name.