Chapter 2

Discourse analysis of games

James Paul Gee

Can there be a field of discourse analysis for games?

A field of discourse analysis applied to video games does not yet really exist. But could there be such a thing? If there was such a thing, what would it teach us about discourse analysis for language?

The question about whether there could be a field devoted to the discourse analysis of games does not ask whether we can analyse games. We can analyse any semiotic system. However, to linguists, discourse analysis builds on syntax and semantics (Gee 2014a). That is, discourse analysis takes as its beginning point 'sentences' or 'utterances' that have already been assigned a structure (syntax) in terms of basic units and their combinations and a semantics in terms of the basic ('literal') meanings of these units and their combinations.

Discourse analysis analyses language in use and it deals with meanings that go beyond semantics and involve context and inference. In my view, discourse analysis studies two closely related things (Gee 2014a). We can call these 'packaging' and 'flow'. First, discourse analysis studies how things are said and written and how they could have been said or written differently and what difference it makes that they were said or written the way they were. For example, why does someone say or write 'It took only an hour for my house to burn down in the fire' versus 'My house took only an hour to burn down in the fire'. In these two sentences information is packaged (combined) in different ways, using the syntactic resources of language with different intentions and expected effects. And, of course, this information could have been packaged into more than one sentence, for example: 'My house burned down in the fire. It took only an hour.'

By the way, it is sometimes argued that there are no sentences in speech (see Gee 2014a for further discussion of this issue). This is, as far as I am concerned, not true. The basic rules of syntax determine what counts as a sentence. Of course, in speech, sentences are often more loosely organised and often more fragmented. Furthermore, intonation plays a major role in what count as units on the order of clauses and sentences.

Second, discourse analysis studies how sentences connect, combine, and pattern across the sequence and flow of time in written or spoken language used in different contexts and situations. For example, why does someone say or write 'My house took only an hour to burn down in the fire' (one sentence) versus 'My house burned down in the fire. It took only an hour' (two sentences) versus 'There was a fire. My house burned down. It took only an hour' (three sentences).

Both in the case of packaging and of sequencing discourse analysis concerns itself not with the sorts of meaning semantics deals with, but with 'situated meaning' (sometimes called 'utterance token meaning'). Situated meanings (Gee 2004, 2014a) are the meanings words, phrases, sentences, and sequences of sentences take on in actual contexts of use. Semantics deals with meaning, or, better put, the meaning ranges (possibilities) of words, phrases, and sentences (this is sometimes called 'utterance type meaning'). For example, at the semantic level, the word 'coffee' means anything to do with the substance coffee. In actual contexts of use the word can have different situated meanings. For example: 'The coffee spilled. Go get a mop' (liquid); 'The coffee spilled. Go get a broom' (grains or beans); 'The coffee spilled. Stack it again' (tins); 'I'll have coffee ice-cream' (a flavour); 'Big Coffee is as bad as Big Oil' (an industry).

Situated meanings are determined by what speakers/writers and listeners/readers take as relevant aspects of context. Situated meanings are also determined by shared cultural knowledge. Such knowledge has been studied under umbrella terms like 'folk theories', 'cultural models', 'figured worlds', 'schemes', 'frames', and others (Gee 2004, 2014a; Holland et al. 1998). Thus, discourse is also related to the study of cultures and social groups that share knowledge and practices with each other.

Thus, the question 'Can there be a discourse analysis of video games?', taken literally, asks whether games have a syntax (a grammar), semantics, packaging, sequence/flow, situated meanings, and associated social and cultural knowledge. If they do, they are in that sense 'like language' and open to discourse analysis.

Before we proceed, let's be clear that to demand that a system have a 'syntax' is to demand that it have basic units that combine in predictable ways into larger units. It is to demand, as well, that the meanings of the larger units be computable in some fashion from the meanings of the smaller units. To demand that a system have a 'semantics' is to demand that its basic units and their combinations have basic meanings or meaning ranges fixed by conventions. While these conventional meanings can vary across different contexts (and new contexts can extend or change their meaning ranges), there must be a conventional core that sets some limits on contextual variation and shapes how contextual variation in meaning operates.

While this chapter presents the view that there could be a discourse analysis for games, the proof would be in the doing, not just the suggesting. We

would actually have to attempt to build the field to see if it could exist and would have any interesting impact. We need to start by considering if and how games have syntax, semantics, packaging, sequence/flow, situated meanings determined (reflectively) by context and social and cultural knowledge

The world has a syntax and semantics for us humans thanks to how human vision works (Marr 1991, 2010). The eye sees the world in vaguely bounded 2D (upside down) images. The eye and brain then process these images in order to construct 3D images with bounded edges and clear shapes. These edges, angles, and bounded surfaces and the way they are combined into spaces and objects (and actions across the flow of time) constitute the syntax of the world for us humans. We then assign names and conceptual labels to the spaces and objects and actions, based on context, cultural knowledge, and social conventions. This is the semantics of the world for us humans.

Scientists have special tools that allow them to see the world in a different way from 'everyday people'. With their telescopes and microscopes, they see different units (like atoms, cells, and stars) that combine in different ways (into molecules, organs, and galaxies). For them, the world has a different syntax and semantics.

Games are made out of a flow of visual images. So they share the syntax and semantics of the human visual world. But, like scientists, gamers have special tools that allow them to see the game world in a different way. Gamers have controllers and avatars through which they can manipulate the game world to accomplish goals and solve problems. Thus, they see the game world not just in terms of spaces, objects, and actions, but in terms of what these things in the game world are good for in terms of accomplishing their goals for winning the game and solving its problems (Gee 2007, 2014b).

Gamers see the game world in terms of what we can call 'game mechanics' (Gee 2007, 2009, 2014b). Game mechanics are what you can do with things in a game. So gamers see the game world in terms of verbs (actions): crates are good for breaking, ledges are good for jumping, shadows are good for hiding, and so forth. Additionally, things in game worlds can combine in various ways to enable certain actions. For example, a ledge, gap, rope, and wall can in some games combine to enable a deft set of moves to get across the game world (as in Tomb Raider games, for example, games in which the famous character Lara Croft dexterously explores caves, ruins, and other mysterious sites).

Though games are built on the syntax and semantics of human vision, their distinctive syntax is composed of the objects and spaces relevant to action in the game. The semantics of a game is a conceptual labelling of these spaces and things not just in terms of their real world identity (e.g. a crate) but in terms of what they are functionally good for in the game (e.g. breakable to get a power-up).

We might say that games have a second-order syntax and semantics based on top of the first-order syntax and semantics of human vision. It is worthy

of note, however, that the syntax and semantics of games, based as they are on 'what actions things are good for' is very much the way we humans look at the 'real' world when we have goals and must take actions we care about to carry them out. There is a real sense in which games 'mimic' our human ways with the world when we are engaged actors (Gee 2014c).

While little work has been done on the discourse of games from the perspective developed here and in Gee (2014c), there are some important related sources. These include: Bogost (2007, 2011); Wolf (2012); Petersen (2012); Paul (2012) and Squire (2011) among others. But, keep in mind, we are not here talking about the language in games or the language gamers use in and out of games. We are talking about games as multimodal forms of digital–human interaction within a system with syntax and semantics and open to discourse analysis in a linguistic sense.

Given that games have a syntax and semantics, we now can ask: Do the ways video games package things, the ways they sequence things, what things mean in actual contexts (situated meaning), and how situated meanings relate to context and culture work, in any significant way, like they do in language? To get at these questions in the small space I have here, I will discuss but one game, *Thomas Was Alone* (*TWA*).

Thomas Was Alone

TWA is a game that uses very simple 2D images. It is about as minimal as a game can get, but for that very reason exposes the basic structure and function of game worlds quite well. Figure 2.1 is an image from *TWA*. Note that while this image is static, the placement of the characters in it was determined by the player. The image is a result of action. Further, the very

Figure 2.1 A screenshot from *Thomas Was Alone*

next screen will be created by the player's actions, as well, based on his or her assessment of the problems to be solved. Each sequence created by the player will reflect, too, the player's interactions with the story and with the emotions of the characters (even though they are shapes!).

If you play the game *TWA*, you immediately see that the small coloured rectangles in the game are like words that can be combined together (like a phrase). For example, players can stack the rectangles on top of each other in certain orders (e.g. red on orange on yellow on blue). Order matters in a minimal way. Each shape has a characteristic movement of its own and each can be moved independently, but when they are combined, the bottom one determines the movement. The bottom one functions like a predicate in language. In the game, the player must get all the shapes to the end of each level, using their different actions and combinations.

So we have a clear, albeit simple, syntax and semantics here. For example, 'red on orange on yellow on blue' means 'stack can move over water' (because that is the blue rectangle's basic action). The order of the stack matters in terms of what subsequent actions are possible, since the shapes can jump off only from the top down.

So how do situated meanings work in *TWA*? Do the shapes and their actions take on specific and extended meanings in actual contexts of use (play)?

TWA has a story. The story is narrated by a narrator whose narration is heard and whose words are printed on the screens. Within the story each shape has a name and something of a backstory as an artificially intelligent agent inside a computer whose programming has gone awry. The shapes are trying to escape the system. In an interesting twist, each shape has certain unique abilities and limitations (determined by the game's game mechanics/semantics) that fit with the character's personality trait and role in the story.

For example, Thomas, the red rectangle, has an up-beat attitude and can do an average jump. John, the yellow rectangle, is arrogant and eager to show off and can jump quite high. Claire, the blue square, who starts off feeling bad about herself but comes to see herself as a superhero, cannot jump well or move fast, but she can float and move in water and thereby save others by giving them rides across water.

The game's story allows us to assign meanings far beyond 'short jumping yellow rectangle'. They allow us to assign emotional and narrative meanings to the rectangles as they act alone and together. The story – and our cultural knowledge about escape stories and about computers – give us cultural models or frames within which we can give richer interpretations to what is happening.

In *TWA* the game's oral narration is also printed on the screen. Since *TWA* prints the words of the narration on the screen, this, in a way, subtracts the words from the oral narration and means that the oral narration mainly

functions to carry the intonation of the narrator's voice, the musical and affective part of speech. This affect is created in part because we can read much more quickly than we can hear, so the player has often read all the (short) material on the screen before the narrator has finished saying it. The player has the 'meaning' but still must pay attention to the intonation contours. The narration in *TWA* is in a British accent that is amazingly good at indicating the emotions of the characters (rectangles though they be), emotions like fear, self-loathing, loneliness, liking and love, caring, arrogance, humility, and trust.

We attribute these emotions and attitudes as deeper meanings for each character by considering the contexts they are in. Consider the image from *TWA* above. Given the words on the screen, the positions of the characters, and the situation we are in in the game at this point, as well as our earlier play in the game, we can attribute to John (the tall yellow rectangle) a situated meaning or inference like: John wants to help, though not necessarily for altruistic reasons, but because he likes to show off and look good to others.

Let's turn now to the ways meanings and inferences are built up in the context of the order, sequence, and flow of screens in the game. Let's assume, for the sake of argument, that a screen prior to the image above had John, the yellow triangle, down on the same level as Thomas (the red rectangle) and Chris (the orange square). Assume further the player has then jumped John – a very good jumper – up to where we see him in the image, up above Thomas and Chris. From this vantage point, a player can clearly see that he or she could move John to the right, down the little alley, and then again further to the right and away from Thomas and Chris.

But the player can also readily see that moving John to the right will not get Thomas and Chris up the ledge so they can move on in the game as well. They cannot alone or together jump high enough to get up the ledge. John must come back down and allow Chris to jump on top of Thomas and then to jump from Thomas to John and, finally, to jump up to the ledge. Thomas can then jump on John and then up to the ledge. And only then can John jump back up by himself. This creates a sequence in the player's mind, a sequence that he or she can then create.

In the image we see that John is higher than Thomas and Chris and that he can easily go on without them. We see, too, that if the game (and its story) is to continue, he must go back down, place himself again on the same level as Thomas and Chris, help them, and then move on last (not first) himself. All these meanings derived from sequence reinforce the sorts of situated meanings we have drawn from the story and contexts of play. John thinks more highly of himself than he does of the others. Forced to go back and help, he has to rationalise this as not a weakness, but as a strength. This strength is not only that his help is essential to mitigate Thomas's and Chris's weaknesses. It is also that John will look good in the act and others will see how special he is.

It is clear from even this simple analysis that how game designer and players (through their choices and actions) create context is a crucial way in which games take on meanings beyond their game mechanics (semantics). Note that in the case of language, context is a co-creation of the world and of how humans construe things in certain ways. As in the case of language, gamers have to know what is relevant (and how it is relevant) in the contexts of their play, which often involve games that are much more graphically complex and rich than *TWA*. Games and players co-create contextual relevance and meaning, as does the world and speakers and writers.

Conversation and affordances

The discussion so far has been meant to be just a mile high overview of a language-like analysis of the structure and meaning of video games. Nonetheless, one very important variable has been left out. Language is used in interactions. We can interact with other people or we can interact with a written text. In either case, there is a sort of conversation going on. So, too, when a player plays a game, the player is having a conversation with the game, indeed a more overt and reciprocal one than readers can have with written texts (which, as Plato long ago pointed out, cannot actually respond to us, see Gee 2011). The player and the game respond to each other in turn.

Although at this point the idea is speculative, it is possible that the notion of 'conversation' could be generalised to cover oral and written language, games (and perhaps other interactive media), and our interactions with the world. I believe that the way to do this would be through a notion well established in ecological psychology, the concept of affordances.

When we humans look at the world in a goal-driven way we actively seek *affordances* in the world. Affordances are what things are good for, based on what a user can do with them (Gibson 1979). For us humans, a hammer is good for pounding nails. That is one of its affordances. A hammer is also pretty good at being a paperweight or a murder weapon. These are others of its affordances. It is very bad at being a toy for infants and you simply cannot use it as food. These are not affordances of a hammer for humans.

Affordances are only affordances, though, given that a potential user of the object has the ability to use the object to carry out the action it affords. The user must have what we can call an *effective ability*, the ability to effect (carry out) the affordance. Humans usually have the effective ability to use hammers for pounding nails. Animals without an opposable thumb do not. They cannot properly hold the hammer. For us humans, hammers do not have an affordance as food. But if they have wooden handles, they do have such an affordance for termites. Termites have the effective ability to eat wood. We humans do not.

Human life and survival is all about finding affordances which one has the effective abilities to put to good use. Let's say you want to get across

a creek. You look around. The log on the ground would afford you the opportunity to cross the creek if you have the ability to move it and good enough balance to walk across it. A line of rocks across the creek would afford you the opportunity to cross the creek, as well, provided you have the ability to walk on wet and perhaps slippery rocks. The creek affords you the opportunity to cross it by swimming across if you can swim across a fast-moving current. Your burly friend could get you across by carrying you if you can convince him to do so and you are able to put up with the humiliation of being carried across like a child.

We look at the world around us to find things with affordances that match our abilities so we can accomplish our goals. Let's call this process of seeking to align or pair affordances with effective abilities the process of 'aligning with the world'. People (and other animals) who are poor at aligning with the world risk danger, failure, and death.

There is a sense in which we humans have conversations with the world, conversations which are formalised in science (Gee 2013). When we form a goal and act on the world, we are looking for the affordances of things in the world, affordances that we have the effective abilities (within the constraints of a given identity) to use to accomplish our goals. Our actions are probes in the world or questions put to the world to see whether and how we can align our effective abilities with affordances of things in the world.

Our conversations with the world go something like this: We have a goal. We take an action in the world, an action that is a type of 'probe' or 'question'. The world responds in some way, answers back. Given that response we ask ourselves if the action led to a result that was good or bad for accomplishing our goal. We 'appreciate' the result of the action in terms of affordances for accomplishing our goal. We then act again and proceed in a probe–response–reflect–probe again cycle until we accomplish our goal, change it, or give up. Of course, things can get more complicated as we pursue more than one goal at a time.

This probe–response–reflect–probe again cycle is a type of conversation with the world. Like all conversations it requires us to listen to and respect our interlocutor (here, the world) if we want to have a good conversation. Science is a formalisation of this sort of conversation with the world we all have. Science has tools for new sorts of probes (questions) and new sorts of reflection on responses. But evidence in science is, at root, always a response from the world.

The probe–response–reflection–probe again cycle is at the heart of video game playing as well. Often via an avatar, gamers form a goal (based on the design of the game and their own desires), act to probe the game world, reflect on the result, see the result as good or bad, and act (probe) again in a chain of acts meant to accomplish their goal. Gamers seek to understand and use the 'rules of the game' to align themselves with the game (in terms

of affordances and effective abilities) properly to succeed, just as scientists seek to understand and use the 'rules of the world'.

In a conversation between two people in language, we have goals we want to accomplish (e.g. bonding, informing, motivating, manipulating, or reassuring our listener or listeners). We probe our listener/s through moves in language (a form of action), reflect on their responses, and then act again based on these responses. In conversations with others, the other is the 'world' we are probing and we are in turn the other's world, since the other has goals as well when they respond to us and take their turn at talk. In conversations with others we seek affordances in their talk, attributes, abilities, desires, skills, character, and language resources for which we have the necessary effective abilities to use (yes, sometimes, manipulate) for our purposes (goals).

So we are arguing that when we humans talk, when we act in the world, whether as part of everyday life or science, and when we play a video game, we are having interactive, responsive, turn-based, conversations based around the search for affordances we can use. Just as Plato thought, reading is a sort of secondary or derivative conversation in which we as readers have to answer for the text (for the implied writer) with respect for that text as a different voice from our own (see Iser 1974 for the notion of implied writers and implied readers).

Note that in this theory, writing, game design, and the design of virtual worlds all involve designing conversational platforms or spaces. And, indeed, it is not for nothing that some scientists think they are coming to understand the 'mind of God' (Davies 1992) when they study the world, since they seek to understand the world's deep design, that is, the deep patterns that allow us to effectively act on and in the world.

The question then becomes: Can we show that conversations in language, interactions with the world in everyday life and in scientific investigations, and video game playing are, at a deep level, similar (though not of course identical)? Better yet, can we learn more about them all by seeing their similarities worked out in different ways in different contexts? If this pursuit turns out to be meaningful, then discourse analysis could be generalised quite far, indeed. We would need to develop more general theories of and tools for conversational interactions where it is not just people that answer back, but games, other media, and the world as well.

References

Bogost, I. (2007) *Persuasive Games: the expressive power of videogames*, Cambridge, MA: MIT Press.

Bogost, I. (2011) *How to do Things with Games*, Minneapolis, MN: University of Minnesota.

Davies, P. (1992) *The Mind of God: the scientific basis for a rational world*, New York: Simon & Schuster.

Gee, J. P. (2004) *Situated Language and Learning: a critique of traditional schooling*, London: Routledge.
Gee, J. P. (2007) *What Video Games Have to Teach Us about Learning and Literacy*, 2nd edn, New York: Palgrave/Macmillan.
Gee, J. P. (2009) 'Playing Metal Gear Solid 4 well: being a good snake', in D. Davidson (ed.) *Well Played 1.0: video games, value and meaning*, 263–274, Pittsburgh, PA: ECT Press.
Gee, J. P. (2011) *Social Linguistics and Literacies: ideology in discourses*, 4th edn, London: Taylor and Francis.
Gee, J. P. (2013) *The Anti-Education Era: creating smarter students through digital learning*, New York: Palgrave/Macmillan.
Gee, J. P. (2014a) *An Introduction to Discourse Analysis: theory and method*, 4th edn, London: Routledge.
Gee, J. P. (2014b) *Good Video Games and Good Learning: collected essays on video games, learning, and literacy*, 2nd edn, New York: Peter Lang.
Gee, J.P. (2014c) *Unified Discourse Analysis: language, reality, virtual worlds and video games*, London: Routledge.
Gibson, J. J. (1979) *The Ecological Approach to Visual Perception*, Boston, MA: Houghton Mifflin.
Holland, D., Skinner, D., Lachicotte, W. and Cain, C. (1998) *Identity and Agency in Cultural Worlds*, Cambridge, MA: Harvard University Press.
Iser, W. (1974) *The Implied Reader: patterns of communication in prose fiction from Bunyan to Beckett*, Baltimore, MD: Johns Hopkins Press.
Marr, D. (1991) *From the Retina to the Neocortex: selected papers by David Marr*, Boston, MA: Birkhauser.
Marr, D. (2010) *Vision: a computational investigation into the human representation and processing of visual* information, Cambridge, MA: MIT Press.
Paul, C. (2012) *Word Play and the Discourse of Video Games: analyzing words, design, and play*, New York: Routledge.
Petersen, J. (2012) *Playing at the World: a history of simulating wars, people, and fantastic adventures from chess to role-playing games*, San Diego, CA: Unreason Press.
Squire, K. (2011) *Video Games and Learning: teaching and participatory culture in the digital age*, New York: Teachers College Press.
Wolf, J. P. (2012) *Building Imaginary Worlds: the theory and history of subcreation*, New York: Routledge.