Design Decisions 2: The Conversation System

Dialog is one of the main ways the player interacts with the world in an RPG, usually following behind killing things and somewhere just ahead of exploration. It is the primary way to learn backstory and flavor, as well as get quests, pointers, and clues.

Picking a conversation system was another of the early choices I had to make in building The Dark Unknown. I knew from the start that the game was going to be tile based, *Ultima V* style in general, but I had a number of options for dialog.

These were the possibilities, as I saw them:

- Single Response: *Ultima III*. This is less a dialog system and more a monologue system, where each NPC just respond to being interacted with with a single response.
- Typed Keyword: *Ultima IV-VI* style. The user types a single word into their dialog box. Further subdivided into *Ultima IV/V* and *Ultima VI* varieties, where in the latter usable keywords are highlighted in the NPC's speech.
- Clicked Keyword: *Ultima VII* style. The user is presented with a list of keywords/phrases and chooses among them. Optionally can include an "Other" option where the user types something not on the list.
- Full sentence dialog: *Infinity Engine/Ultima IX* style. Player is presented with full sentences, possibly voiced.
- Dialog Wheel: *Mass Effect/SWTOR*. What you select shows what kind of response you'll give, and you don't even know what you're actually saying until you say it! (*SWTOR* is less extreme in how surprising what you actually say might be.)

When I was first starting to work on the *Dark Unknown* engine, I put up a poll to see what people would prefer to see. (On LiveJournal, to let you know how long ago this was!) At that point, and with that audience, it was a pretty close split between Typed and Clicked Keyword, with the addendum that there was near universal preference for making sure the keywords were highlighted if I went with the typed keyword system.

In the end, I went with typed keyword over clicked for two primary reasons—first, while typed keyword has appeared in some games with character portraits (*Ultima VI*) and some without (IV and V), clicked keyword has nearly always been presented with portraits. And so, as I did not think I had the budget for character portraits, the decision was made. This then reinforced my decision that the game should be entirely playable without a mouse—had I had a preference for a clicked option, I would have either come up with a UI that allowed you to select with the keyboard, or rejiggered the UI to be keyboard/mouse, but since I had other reasons to want to stick with keywords I also stuck with the keyboard.

But what are the pros and cons of each? Let's go through some of them. While I will define some of my terms in more detail after I go through each type, let me give you a few concepts before I get into the types of conversation systems:

- **Nuance** is the ability to show nonverbal cues that would accompany the dialog.

- **Flexibility** is a measure of how much freedom the player has in how they pursue the discussion.
- **Responsiveness** is the ability for the dialog, the responses, and the options to vary based on changes in the game's state. There are two kinds of responsiveness- whether the NPC's responses can change based on what is happening in the game, and whether the player's options for prompts can reflect the situation. The former is trivial- even the "single response" style of conversation can have responses that vary, so long as the engine supports it. But the latter is more interesting, and it varies by conversation style.

Single Response

This is obviously the most restrictive. Pros include it being a lot faster to write, and if your game is one where the emphasis is on combat, say, this minimalist conversation system will stay out of the way and let the player get to what they are there for. The primary con, of course, is its extreme shallowness. This type of system is neither **flexible** nor **responsive**, and not only can it not show **nuance**, it is devoid of room for nuance.

Typed Keyword

Seen as somewhat primitive these days, now that pretty much all modern big-budget games go with full voice acted sentences, but there can be surprising depth even in these systems. Among its pros are that it doesn't put words into the player's mouth, allowing the player to identify as much as they wish with the character they are playing, and maximal **flexibility**. Cons are that it is nearly impossible to show **nuance** in responses with any subtlety, and that it is nearly impossible for dialog options to be **responsive**.

Also, there are memory issues- as each new response needs to be typed by the player they need to remember each thing that they wish to ask or talk about. Depending on the type of gamer, this can either be a pro or a con, depending on your view of requiring players to take notes during their playing. There is a real con here, though- unless care is taken to ensure that the player can always recreate dialog, it is possible for someone to step away from the game for a few months and return not only unable to remember what was going on, but be unable to figure it out. (For example- if you have a monster whose last words as it dies include a keyword you must ask the local lord about, if you don't write that down it is lost forever, as the monster is dead.)

There is, of course, nothing that prevents an in-game journal/quest log/chat transcription being available from within a game that uses a typed keyword system and wishes to have certain text only appear once.

I will admit that I went into *The Dark Unknown* assuming that this would be substantially *easier* than full sentence dialog. Fewer words to write, at the least. But it does have its own challenges- it can be difficult to structure the reveal of information or NPC responses around single word prompts.

One that I had not foreseen was what prompted me to include **responsiveness** in my list of qualities. It is also very difficult for this system to allow certain kinds of player character

reactions. For example, if you walk into a room and are surprised to see someone there, there's no way to give you the option to say, "What are you doing here?" None of those individual words clearly convey the concept- it would require a late Infocom-style parser to allow it to be entered by the player at a prompt. One of the systems below, however, which present the player with a list of options to choose from, can, when you enter the room, give you a set of options like, "Say 'What are you doing here?' OR Enter silently."

Fundamentally, this means that if you are using this system, as I am, you need to design your dialog and your encounters to not highlight these deficiencies.

Clicked Keyword

Moving closer to the modern era, the clicked keyword system took over basically at the same time as the mouse did for computer gamers, while console controllers were always much more strongly in alignment with them. Some pros: like typed keywords, doesn't really put words in your mouth, leaving you to read between the lines to imagine how you are saying or asking whatever you've just said. The primary pro is that it removes much of the need to take notes on conversations, as when Biff tells you to ask Janet about GRUES, the keyword is right there on screen to remind you when you next speak to Janet. (You may still want to take notes to remind yourself that you need to talk to Janet, though.) Also, the clicked keyword system allows keywords to be longer than one word- while this was not technically impossible with typed keywords, it is very impractical. **Nuance** becomes possible, though was rarely implemented, as it has a better mental model for the player when full sentences are involved. **Responsiveness** similarly becomes possible. The primary con, then, is the loss of **flexibility** unless care is taken to provide a way to allow it in the dialog UI. Other cons include the possibility of overcrowded screens, if there are a lot of things to go through. This, again, can be mitigated with care by for instance grouping inputs by subject. Some clicked keyword games might just always list *all* of the topics you can ask about, while others might first present you with something like this:

> *What would you like to ask about?*

- *You*
- *Monsters*
- *Treasure*

Selecting "You" might bring up a new list:

> *Ask what?*

- *Name*
- *Job*
- *Parents*
- *…enough about this*

And selecting "enough about this" returns you to the top level. This has the potential downside of making the player guess which category something is in. More on this kind of thing later.

Full Sentences

This option may include full voice acting, either for the entire game (most modern games such as *Skyrim*), or only important NPCs (*Persona 3-4*), or only important lines (*Baldur's Gate 2*). In modern games voicing is assumed, but it is by no means a requirement. This has many of the same pros and cons of clicked keywords- once again there are fewer memory issues, as the response will just be there when you talk to the NPC. Perhaps the primary con is that it puts words in the mouth of the player- if I type "paladin" into a keyword system I could imagine that I am saying "what can you tell me about the local paladins' guild?" while here I might only have the option to say "What's a paladin?" It also increases the load on dialog writers, as your responses and prompts need to be longer. (As I noted earlier, this is not as much of an increase as I initially had expected, as it turns out creating dialog beyond a certain minimal complexity with the keyword system is its own kind of challenge!) Full sentences can often have **nuance** and be **responsive** but we have fully lost the **flexibility** that simpler systems can have.

Conversation Wheel

This is a newer option which I am not sure had yet been developed when I started my project, but I include it for completeness now. In this, you are presented with a general idea of what you will say without it being fully written out for you. The wheel allows for a considerable amount of **nuance**- you can always arrange it so an option that appears in the upper left is the "lawful" option while the lower right is "chaotic", perhaps, or have the wheel itself show colored wedged to go with the options. In addition, this gives it the ability to do a good job of showing what the tone of your response will be. This has most of the pros of the full conversation system above. As a con it is even worse about putting words into your mouth, as not only do you not get to choose the words, you might select an option and discover that the full version was *not* something you had meant to say! Like full sentences, it is difficult to be **flexible** and easy to be **responsive** in this system.

Voice Acting

VA can be a part of several kinds of dialog system, though most commonly full sentence and conversation wheel. It adds an impressive level of immersion, but is expensive, and that expense means you really don't want to make changes to someone's dialog and need to rerecord it. With text, a last-minute change is doable, but with VA it's much more dicey.

Character Portraits

This is arguably a bit of a digression, but while portraits are not exclusively tied to dialog, there's a pretty strong bond there. You can have character portraits with any of the styles of dialog- Ultima VI had them with typed keyword, VII with clicked keyword, Baldur's Gate with full sentences. The only reason there aren't many examples of portraits with a conversation wheel is that once 3d got good enough, portraits were replaced with simply zooming in on whomever you were speaking to.

There are a couple things character portraits do very well. The first and arguably most important is it makes it easier for the player to keep track of who they are talking to, and

recognize them again later. It is so much easier to recognize a picture than to keep straight a name and small in-game representation that may not even be unique.

The second thing is adding nuance and emotion to the dialog. Not all games with portraits had enough of them to do so- Ultima VI still used text to emote if someone looked nervous, or pleased, or was blushing, or what have you. But you have the option of multiple portraits for the same person, as you'll see particularly in JRPGs- one blushing, one sweating, one angry, etc. If you don't want multiples for everyone, you could have two or three for just a handful of important people- Batlin in Ultima VII is one example there.

The last thing is that they make it easier to have a third party interject into a conversation. You can do it with just text, but showing the picture makes it easier to track when someone new is speaking, and also helps you remember who the other person is, as previously discussed.

Defining Terms

Now it is time for me to go into more detail about the bolded terms above: "nuance", "flexibility", and "responsiveness". These aren't industry standard terms or anything, but they're what I've come up with to describe things I look for in dialog.

Nuance is the ability to know what the subtext of what you're saying will be. My favorite example came in *Planescape: Torment* when at one point you are given a set of options where two are identical, but one of them adds "[Lie]" to the option. I include in this when a game tells you that your choice is "renegade", "paragon", "dark side", or what have you, and also when a response involves a die roll (a Persuade check, perhaps, or maybe Intimidate) that will have consequences when failed. Nuance is difficult to impossible in systems that rely on the player to enter their own inputs for dialog- it has some tension with the reading between the lines that is available in systems where you are choosing a topic or just a keyword and it is left to the player to think about what the character is actually saying, and comes into its own when the game presents you with a full sentence, where it can add to your understanding of your response by letting you know that you're lying, being sarcastic, or offering a hug.

Flexibility covers a number of bases. One element is that you allow the player to guide the conversation to a greater extent- rather than giving a usually small list of responses to the current moment of dialog, the player can say anything they want. This can cause a lack of immersion (if you think about it too hard you might wonder why no one minds that you jump around from topic to topic all the time! But I find it worth noting that you already have to either allow the immersion breaking of being able to make NPCs repeat themselves, or you have to have (a) a good journal system and (b) a lot of NPCs that either run out of things to say or have infinite dialog), but it allows a lot of freedom. If the player has made a connection that the game doesn't necessarily know you've made yet, it allows you to ask about it- an example here might be drawn from *Ultima IV*. Once you learn that you will need a Rune and a Mantra from each town, you can just walk up to every NPC and ask them about it, rather than needing to find the one that mentions it on their own. This can also allow for easter eggs- responses that nothing has pointed to, and are not (necessarily) game important. This can do an excellent job of fleshing out your game world, but doing so requires putting work into writing dialog that most players are unlikely to see. Find the right balance wisely. (I'm hoping I will.)

Flexibility is largely a hallmark of typed/keyword conversation systems, though it can be obtained in a system that presents you with options to choose from by (a) grouping responses into categories, and (b) adding an "Other" response option that then allows you to type something in. *Ultima* by and large didn't do this, with the non-conversational exception of *Ultima III*'s "(O)ther" command, which allowed you to do things like Bribe guards without it being obvious that bribery is among the things that you can do (as it would be if it just appeared in the list of keyboard commands). Remember also that even if you provide an "Other" command, people are less likely to make use of it (unless it is necessary to progress) than they would be to try typing in interesting things while still in the primary mode of the dialog interface, as they would be in a keyword system.

One consequence of high flexibility is that it allows sequence breaking. Replay *Ultima V* (or play with a walkthrough at hand) and you can go straight to the last person in any conversation chain, walk right up to the Resistance leader and give the password without having ever spoken with anyone who might have told it to you. You can either embrace this (is it really so bad if someone skips ahead if they know what they're doing?) or guard against it (*The Dark Unknown* makes heavy use of game flags to monitor whether you're at this step of some of the quests, though it allows sequence breaking in others).

Responsiveness is what I term the game responding to changes in your character and the game state. An example from my own experience- at one point in *The Dark Unknown*, someone tells a long rambling story to which there were no witnesses. I would have loved to give you an additional possible response, something like "how do we know that?", but only if you had a sufficiently high intelligence. But as I don't give you prompts to choose from, it was not possible and so I just had the NPC volunteer that there's no way to know if it's true but it makes a good story.

One level of responsiveness I *could* have had would have been to have that NPC only volunteer that last few lines if your intelligence is high enough- that's also responsiveness, from the other side. I consider it less interesting because the player doesn't get to make a choice about how to bring up what they have noticed as a result of their improving their intelligence.

It is easier for games that present you with choices for dialog to be highly responsive, as it is doing a more direct job in guiding the narrative, by not only presenting the NPC's dialog but also presenting you with your choices in responding. A highly responsive dialog system moves things forward by always presenting the player with options, at the possible expense of spoiling things for a player who might not have put two and two together yet only to have the option to tell someone that it was four.

This brings us to another point, which is the **role of dialog** in your game. If you are making a game heavy on combat, or puzzles, then dialog should be limited to just what is necessary to get the player from set piece to set piece and otherwise getting out of the way of the player. But if you are making a detective mystery, talking to people will be one of the key mechanics and the dialog should be lovingly tended into something that deeply affects the player. Take a look at what things allow a player to *progress* in your game- how many are killing a boss? Finding a treasure? And how many are learning from an NPC, or solving an NPC's problem? It's

important to make sure that your dialog is integrated into the rest of the game, rather than feeling like a chore to be spacebar-ed through to get to the good parts.

Bringing it all together

In the end, my decision was made. Partially due to budget—I couldn't afford character portraits—and partially due to feeling it was more in keeping with the style that the rest of the game would adhere to, I chose to use the **typed keywords** style in The Dark Unknown. I am making as much use as I can of the extra flexibility this affords me- making my dialog trees as wide as possible, giving each person a handful of things to respond to that are not directly game related, and are not pointed to by any other NPC, but are things it would just make sense for them to have something to say about. Names of their friends, the name of their town, or current events, as examples. I'm using game flags to prevent jumping ahead in places where it makes sense, but am allowing it in other places. I do have third parties jumping in on conversations sometimes, but I try to make sure it is always clear who is speaking. I have some further experiments in mind for how to make that clearer in the UI. There are things I didn't realize going in- as I noted earlier, I really felt like typed keyword was going to be easier than full sentence. And I haven't tried full sentence responses yet, maybe it is! But typed keyword has been a challenge in a lot of ways, mostly in finding a natural way for information to be presented to the player.

I hope you found this interesting. I think next time I may talk about how I represent dialog in the backend- my conversation editor and how I store it in the data. But that is for another time. Tune in then!