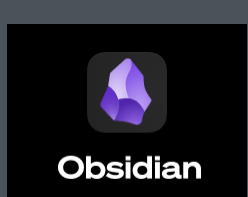


By JOHN GRUBER

- ARCHIVE
- THE TALK SHOW
- DITHERING
- PROJECTS
- CONTACT
- COLOPHON
- FEEDS / SOCIAL
- SPONSORSHIP



Obsidian: the private and flexible writing app that adapts to the way you think. Sign up by Jan 1st for a special offer.

Google Demos Duplex

Wednesday, 27 June 2018

Google has finally done what they should've done initially: let a group of journalists (two groups actually, one on each coast) actually listen to and participate in live Duplex calls.

Heather Kelly, writing for CNN:

For one minute and ten seconds on Tuesday, I worked in a trendy hummus shop and took a reservation from a guy who punctuated his sentences with “awesome” and “um.”

“Hi, I’m calling to make a reservation,” the caller said, sounding a lot like a stereotypical California surfer. Then he came clean: “I’m Google’s automated booking service, so I’ll record the call. Um, can I book a table for Saturday?”

The guy was Google Duplex, the AI-assisted assistant that made a stir in May when CEO Sundar Pichai unveiled it at its Google I/O developer conference. That demo, shown in a slick video, was so impressive that some people said it had to be fake.

Not so, says Google, which invited clusters of reporters to Oren’s Hummus Shop near its campus in Mountain View, for a hands-on demonstration. Each of us got to field an automated call and test the system’s limits.

But, regarding the [curious recordings](#) played on stage at I/O in early May:

Scott Huffman, the VP of engineering for Google Assistant, conceded that the demo at I/O in May “maybe made it look a little too polished.” That’s because Pichai tends to focus on Google’s grand visions for the future, Huffman said.

Ron Amadeo, writing for Ars Technica:

Unfortunately, Google would not let us record the live interactions this week, but it did provide a video we’ve embedded below. The robo call in the video is, honestly, perfectly representative of what we experienced. But to allay some of the skepticism out there, let’s first outline the specifics of how this demo was set up along with what worked and what didn’t. [...]

During the demonstration period, things went much more according to plan. Over the course of the event, we heard several calls, start to finish, handed over a live phone system. To start, a Google rep went around the room and took reservation requirements from the group, things like “What time should the reservation be for?” or “How many people?” Our requirements were punched into a computer, and the phone soon rang. Journalists — err, restaurant employees — could dictate the direction of the call however they so choose. Some put in an effort to confuse Duplex and throw it some curveballs, but this AI worked flawlessly within the very limited scope of a restaurant reservation.

Here’s the video Google has provided. It is indeed an impressive approximation of a human speaking. One thing that stands out, in fact, is the difference between the artificial voice of the Google Assistant on the woman’s phone — no um’s, no ah’s, robotically precise — and the decidedly un-robotic voice of Duplex on the phone call.

Regarding the actual rollout to actual users, some unspecified number of “trusted testers” will get access to Duplex very soon, but only for asking about restaurant hours, not making reservations — and the haircut appointment feature has no delivery date other than “later” and wasn’t demonstrated to the media.

Dieter Bohn, writing at The Verge:

If you’re hoping that means you’ll be able to try it yourself, sorry: Google is starting with “a set of trusted tester users,” according to Nick Fox, VP of product and design for the Google Assistant. It will also be limited to businesses that Google has partnered with rather than any old restaurant.

The rollout will be phased, in other words. First up will be calls about holiday hours, then restaurant reservations will come later this summer, and then finally hair cut appointments will be last. Those are the only three domains that Google has trained Duplex on.

Bohn on the speech quality:

The more natural, human-sounding voice wasn’t there in the very first prototypes that Google built (amusingly, they worked by setting a literal handset on the speaker on a laptop). According to VP of engineering for the Google Assistant Scott Huffman, “It didn’t work. we got a lot of hangups, we got a lot of incompleteness of the task. People didn’t deal well with how unnatural it sounded.”

Part of making it sound natural enough to not trigger an aural sense of the uncanny valley was adding those ums and ahs, which Huffman identified as “speech disfluencies.” He emphasized that they weren’t there to trick anybody, but because those vocal tics “play a key part in progressing a conversation between humans.” He says it came from a well-known branch of linguistics called “pragmatics,” which encompasses all the non-word communications that happen in human speech: the ums, the ahs, the hand gestures, etc.

I’m on the fence regarding the issue of whether it is ethical for Duplex to speak in a way that sounds so human-like that the person on the other end of the call might never realize they’re speaking to a bot. What raises a flag are the injected imperfections. If they’re good for Duplex to use while making a call, why doesn’t Google Assistant speak similarly when you, the user, *know* you’re talking to a bot?

The fact that they started getting fewer hangups when they added these natural-sounding imperfections makes sense. But it’s disingenuous to say they’re not using these um’s and ah’s to trick the person into thinking it’s a human. That’s exactly what they’re doing. The problem is, *tricking* sounds devious. I’m not sure it is in this case. It’s just making the person on the call more comfortable. We use “tricks” in all of our technology. Motion pictures, to name one example, don’t actually move — they’re just a series of still images played quickly enough to fool our eyes into seeing motion.

With or without Duplex’s involvement, the restaurant is going to get a phone call for the reservation. (Duplex doesn’t make phone calls for restaurants that support online booking through OpenTable — at least not if the device user has an OpenTable account.) Based on these examples, Duplex doesn’t seem to waste the restaurant’s time — the phone calls take about the same time as they would if you, the human, made the call yourself. So neither the restaurant nor the employee who answers the phone lose anything when a call is made by Duplex, whether they realize they’re talking to an AI or not. No one is getting cheated, as in the case with, say, bots that play online poker.

To me, the truly difficult ethical questions are years down the road, when these AI’s get close to passing an open-ended Turing test.

Lauren Goode, writing at Wired:

I then asked whether there were any allergies in the group. “OK, so, 7:30,” the bot said. “No, I can fit you in at 7:45,” I said. The bot was confused. “7:30,” it said again. I also asked whether they would need a high chair for any small children. Another voice eventually interjected, and completed the reservation.

I hung up the phone feeling somewhat triumphant; my stint in college as a host at a brew house had paid off, and I had asked a series of questions that a bot, even a good one, couldn’t answer. It was a win for humans. “In that case, the operator that completed the call — that wasn’t a human, right?” I asked Nygaard. No, she said. That was a human who took over the call. I was stunned; in the end, I was still a human who couldn’t differentiate between a voice powered by silicon and one born of flesh and blood.

It’s a shame that Google wouldn’t release the recordings of the calls the journalists answered. Goode’s anecdote above, to me, is the most fascinating of the bunch, and I’d love to hear it. She was able to trip up the logic of Duplex by asking about allergies and high chairs, but was unable to discern when an actual human took over the call. Google’s breakthrough isn’t how smart Duplex is, but how human-like it sounds.

I still think the whole thing feels like a demo of a technology (the human-like speech), not a product. Google claimed this week that Duplex currently succeeds 4 out of 5 times at placing a reservation *without* a human operator’s intervention. That’s a good batting average for a demo, but untenable for a shipping product at Google’s scale. With a 20 percent failure rate, Google would need an army of human operators standing by all day long, to support a feature they don’t make any money from. I’m skeptical that this will ever be a product expanded to wide use, and if it is, it might be *years* away. Google said as much to Ars Technica:

“We’re actually quite a long way from launch, that’s the key thing to understand,” Fox explained at the meeting. “This is super-early technology, somewhere between technology demo and product. We’re talking about this way earlier than we typically talk about products.”

Right now it feels like a feature in search of a product, but they pitched it as an imminent product at I/O because it made for a stunning demo. (It remains the only thing announced at I/O that anyone is talking about.) If what Google really wanted was just for Google Assistant to be able to make restaurant reservations, they’d be better off building an OpenTable competitor and giving it away to all these small businesses that don’t yet offer online reservations. I’m not holding my breath for Duplex ever to allow anyone to make a reservation at any establishment.

- PREVIOUS: [10 Strikes and You’re Out — the iOS Feature You’re Probably Not Using But Should](#)
- NEXT: [Apple Updates MacBook Pro Models With Touch Bar](#)