A Tongue Input Device for Creating Conversations

Ronit Slyper, Jill Lehman, Jodi Forlizzi, Jessica Hodgins
Disney Research Pittsburgh and Carnegie Mellon University
Pittsburgh, PA
{rys,jef,forlizzi,jkh}@cs.cmu.edu

Figure 1: Our tongue joystick allows an actor to invisibly select prerecorded audio clips, letting the articulated-head dragon converse with children in its own voice.

ABSTRACT
We present a new tongue input device, the tongue joystick, for use by an actor inside an articulated-head character costume. Using our device, the actor can maneuver through a dialogue tree, selecting clips of prerecorded audio to hold a conversation in the voice of the character. The device is constructed of silicone sewn with conductive thread, a unique method for creating rugged, soft, low-actuation force devices. This method has application for entertainment and assistive technology. We compare our device against other portable mouth input devices, showing it to be the fastest and most accurate in tasks mimicking our target application. Finally, we show early results of an actor inside an articulated-head costume using the tongue joystick to interact with a child.

ACM Classification: H5.2 [Information interfaces and presentation]: User Interfaces. - Input devices and strategies.

General terms: Design, Human Factors

Keywords: mouth, interface, dialogue tree, turn-taking.

INTRODUCTION
The Disney theme parks have the goal of bringing characters to life by having them act out and describe their stories. Many of the costumed characters in the parks, however, cannot converse with guests. Actors in these costumes wear large fur or plastic head pieces that prevent them from talking. Several of the characters also have highly recognizable accents and speech patterns, which actors cannot replicate accurately. These characters are reduced to using body language and gestures to communicate.

UIST'11, October 16-19, 2011, Santa Barbara, CA, USA.

Articulation abilities have recently been added to some of these costumes. Motors in the eyes and mouth let them open and shut, as shown in Figure 1. This technology is currently used in shows in the parks, where the mouth moves in sync with prerecorded audio coming out of loudspeakers.

We would like to use this articulation technology to allow these characters to converse directly with guests. Because these characters have such distinct voices, the actors inside the costumes will need to trigger context-appropriate prerecorded audio snippets. This situation suggests using a dialogue tree and a heads-up display inside the costume to show the current location in the tree. Although a dialogue tree cannot cover the full scope of conversation, the theme parks present a constrained situation. Actors for the current talking characters in the park (e.g., princesses) tell us that conversations with guests tend to be character-driven: children frequently turn shy and let the actor lead the conversation; and when children actively interact, they ask the same questions (e.g., princesses are always asked where their prince is).

In this paper, we are primarily concerned with the input device that the actor inside the costume will use to navigate a dialogue tree with a branching factor of four. We consider four to be the minimum number of responses to hold a reasonable conversation: if a character asks a question, the guest's response can be categorized as 1) affirmative, 2) negative, 3) other (question-dependent), or 4) uncooperative; the character should have a response ready for any of these four cases.

Common input devices are not usable in this selection task, because they rely on the hands. A character's hands are visible at all times, waving, gesturing, and signing autographs. Instead, we seek an input method using the mouth. We need a device that has a low error rate and is fast enough to enable smooth conversations.

We compare three mouth input devices, using breath, teeth, and tongue. For breath, we use the leading portable mouth input device for people with quadriplegia, the sip/puff switch.

Device	Direction			
	Left	Right	Up	Down
Sip/puff switch	soft puff	hard puff	soft sip	hard sip
Bite senor	two short bites	short long	long short	long long
Tongue joystick	press left	press right	press up	press down

Figure 2: All three devices, ready for testing. From left to right: sip/puff mouthpiece, bite sensor, and tongue joystick. Four-direction selection is accomplished using the actions in the table.

For the teeth, we built a binary bite sensor. For the tongue, we built a new device. Its construction demonstrates using rapid prototyping and an unconventional combination of silicone and conductive thread as a flexible method for creating cheap, low-actuation-force devices.

We tested these devices with a user study that mimics aspects of our target application. The study included a speed test and a conversational interaction using a four-branch dialogue tree. Figure 2 summarizes how each device was used to select among the four directions. We found the tongue joystick to be both the fastest and most accurate. Dialogue latencies when using the tongue joystick ranged from 3 to 4 seconds, which is fast enough to hold a conversation with a child [6]. No significant difference was found between the bite sensor and the sip/puff switch.

Our results have implications for the study of physical and situational impairments [27]. We show that the tongue is agile enough to select among multiple discrete locations, and thus should be further explored as an input device for people with quadriplegia or who otherwise cannot use their hands. Although preliminary work has been done in the area of capturing tongue gestures [26, 12], we are the first to test a form factor that requires no customization or attachment to the user.

In the following section, we describe less common input methods, and show why we converged on the *mouth* as the best high-degree-of-freedom input modality for this situation. Next we describe the three mouth devices. We detail the protocol of our user study to compare them, and show the results of the study. Based on the results, we select the tongue joystick to incorporate in a functional prototype system for an articulated character head. We conclude with possibilities for future work, both in our application and in assistive technology.

RELATED WORK

Nearly all of our computer interfaces today – keyboard, mouse, touchscreen, driving wheel, button pad, gamepad – rely on fine dexterity in the hands. In our application, however, the hands cannot be used for input. An actor's hands must be free to interact with guests.

The human voice also has incredible degrees of freedom. Unfortunately, we found that sound transmits easily through our character head, ruling out standard voice recognition or nonverbal command inputs such as vowel [3] or duration of sound [13]. Whisper recognition or throat-microphone input would be possible, but to achieve high levels of accuracy, command phrases would need to be several syllables in length and easily differentiable, e.g. "show me A", "select option B", inducing a confusing lack of parallelism as well as slowing down the system.

Interfaces developed for people with physical disabilities show that other parts of the body can be used to communicate with computers. Any feature that can be tracked in two dimensions can serve as a substitute for the mouse. Head tracking can be done with an infrared camera [21] or with an inexpensive standard webcam [2, 18]. Eye tracking, using infrared reflection [8] or electrooculography [17, 7], requires less movement, but any type of eye-based input will monopolize the actor's attention. This compromise is acceptable for assistive technology, but not for our application, where the actor must attend to the child.

Any high-level voluntary movements that can be captured can be used for switch input. Examples include foot switches, eyebrow switches, and blink switches. These would be awkward or tiring to accomplish in costume.

Low-level signals from the body also present interesting opportunities. Electromyography (EMG) has been used to control artificial limbs [16]; attempts have been made to use it for mouse input [14] and wheelchair control [10]. Electrical skin potential has been tested as a switch input, but provides low accuracy [19, 20]. Brain-computer interfaces continue to make progress, although they lack the speed and accuracy required for our application. Current research efforts involve techniques such as fMRI, electroencephalography (EEG), and electrocorticography (ECoG); see [11] and [29] for good overviews.

The mouth provides numerous methods of input; some of these can even provide direct spatial mapping. Sip/puff interfaces provide a single degree of freedom, but are portable and easy to use. BLUI [24] is an interface that extends the idea of using blowing as a means of input, but adds a spatial dimension, using a microphone to plot the location of a puff on a laptop screen. Breath can also be spatio-located using thermotransducers [9] and piezo film sensors [15]; the latter work adds selection using tooth-touch recognition with a bone-conduction microphone.

Non-portable mouth devices commonly in use include joysticks, manipulated with the chin, lips, or tongue [5], and mouth sticks, held in the mouth and used to tap directly on a keyboard or touch screen.

Our work was inspired by Huo and Ghovanloo's Tongue Drive System [12]. The system consists of headgear containing magnetic sensors, and a small magnet affixed to the user's tongue. Voluntary motions of the tongue are classified and translated into powered-wheelchair control commands. We were drawn to this system for several reasons. First, as Huo and Ghovanloo point out, tongue muscle is similar to heart muscle, not tiring easily. In our application, we must be wary of repetitive stress injuries. Second, the tongue is very fast. In a similar setup, Struijk [1] constructed a retainer with embedded coils inductively triggered by the proximity of a magnet attached to the tongue. Due to the stringent safety requirements of our application, we sought a self-contained device without the risk of swallowing a magnet.

Recent work by Saponas et al. [26] uses optical infrared sensors to recognize four tongue gestures. The proximity sensors are placed in the right, left, front, and back of a dental retainer. Simple heuristics based on the pattern and timing of triggering are used to recognize one of four tongue gestures: left swipe, right swipe, tap up, and hold up. The 92% accuracy subjects achieved shows this approach to be promising. The Tongue-touch keypad [22], a custom-molded dental plate containing nine buttons, was briefly on the market. The Tonguepoint [25], a mouthpiece containing an isometric joystick, showed that tongue dexterity improves with practice. These devices all require a custom retainer to achieve a low profile, but still present reasonable alternatives to our device.

DEVICES

Our target application, real-time conversational interaction, created a set of requirements for our input device:

Speed: The device must be fast and impose low cognitive load, as the actor will be simultaneously acting and conversing.

Accuracy: Wrong selections would either trigger non sequiters or slow down a conversation as the actor navigates backwards in the dialogue tree; the device must have near-perfect accuracy.

Portability: The device must be self-contained and fit inside the costume.

User independence: For convenient and low-cost testing and replacement, the device will preferably not be customized to fit a particular actor.

Ruggedness and safety: Character costumes are in use day in and day out; the device must be designed without weak points.

We explore which of the three input modalities of the mouth – the breath, teeth, and tongue – can be used to best meet these requirements. We represent these three modalities by portable input devices that attempt to make optimal use of their affordances: a sip/puff switch (breath), bite sensor (teeth), and tongue joystick (tongue). We present details of the devices in the following three sections.

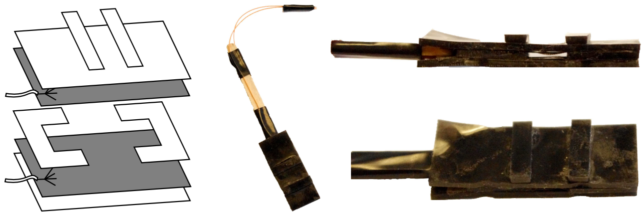

Figure 4: *Left*, an exploded view of the bite sensor's construction; white layers are insulating rubber, grey layers are conductive tape with attached wires. *Center*, the finished bite sensor with mounting stick. *Right*, side and top view.

Sip/puff switch

We used an off-the-shelf USB sip/puff switch from Origin Instruments Corporation [23]. Each subject in our study was given a new filtered mouthpiece when testing the device.

Four-direction selection with the device was done using soft puff, hard puff, soft sip, and hard sip to represent left, right, up, and down, respectively. This combination of breath force and direction is a common way to get four commands. Only the maximum (or minimum) value of the force signal was relevant, not the duration. Softness and hardness were assigned using pre-determined thresholds. Signals less than 80 milliseconds long were dropped as transient.

Bite sensor

Commercial binary bite switches are used for skydiving photography and in cases of severe disability, e.g. ventilator dependence. These switches cost around $100 [4]. Figure 4 shows the construction of our lower-cost switch. We use $\frac{1}{16}$"-thick polyurethane rubber, Shore 60A, and conductive adhesive tape. Thin top rectangles let the user locate the bite area by feel. For each study participant, the bite sensor was covered in a new piece of thin nitrile (a finger cut off an examination glove).

For four-direction control on this one-bit device, we used bite duration to add a second bit. The four directions were represented by the four combinations of short (< 200 msec) or long (≥ 200 msec) bite followed quickly (in < 500 msec) by another short or long bite.

Tongue joystick

We designed a new device for the tongue that captures our goals of portability and user independence. Although we could have used any number of sensing technologies, e.g. camera-based tracking, or the magnetic or infrared sensors from previous tongue work, we chose binary switches because they require minimal processing time and encourage the user to be speedy by provide a tactile "click" as soon as a gesture has been accomplished.

For the form factor, we looked for inspiration to other objects that we manipulate in our mouths: we chew on gum, lick lollipops, bite pencil erasers and worry pen clips; and, as babies,

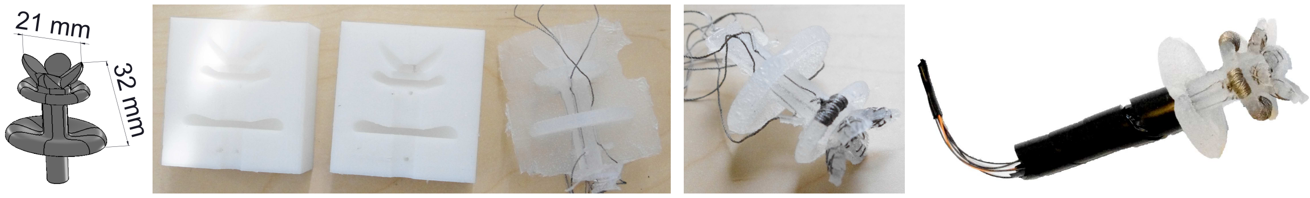

Figure 3: The tongue joystick molding process. From left to right: 3D model; tongue joystick coming out of the printed mold; contact pads being sewn out of the embedded conductive thread; the finished product.

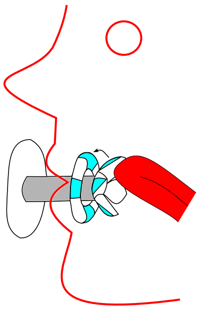

Figure 5: The user bites on the stem (grey), then pushes one of the petals with his tongue (red). The petal contains a contact pad (cyan) which touches the ring beneath, also containing a contact pad (cyan).

we suck on pacifiers. The pacifier turned out to be the perfect inspirational form factor – it is designed to lodge itself securely in a baby's mouth, with the baby's cheeks and gums braced in the stem between the bulbous tip and cheek guard.

Our device is shaped like a pacifier, but with the bulbous tip turned into a flower shape with a ring and four "petals". The tongue can easily align itself perpendicularly to a petal to push on it, as shown in Figure 5. (An early test using a four-directional joystick inserted in a pacifier had shown that though the tongue is dexterous, it actually cannot exert much sideways force.)

Figure 3 shows the steps in the construction of the tongue joystick. We designed the device in a 3D modeling program, and 3D-printed a plastic mold. To build the device, we fixed four strands of conductive thread to the inside edges of the mold, and one going through the center. We then poured Shore 40A RTV silicone into the mold. The device was cured and demolded. Then for each strand protruding from the silicone, we threaded it onto a needle, and "sewed" contact pads using the silicone as "fabric". The central conductive thread was sewn to the undersides of the top petals, forming the ground layer for the contact switches. The four outside threads were each sewn to a segment of ring beneath a petal, creating the signal layer.

This combination of materials – silicone and e-textile's conductive thread – creates a device which is completely soft and basically unbreakable. The silicone switches close with a low enough actuation pressure that they can be comfortably pressed by the tongue. Typical low-actuation pushbutton components close around 150 gram force; our switches close around 50 gram force.

Four-direction control is achieved through direct spatial mapping. In the study, signals less than 50 milliseconds long were dropped as transient. We covered the device in a new piece of plastic wrap for each subject in the user study; in production, we imagine the device having a removable thin stretchy silicone skin in the shape of its convex hull.

USER STUDY

We evaluate these three devices with tasks involving four-directional selection. In this setting, these devices represent different tradeoffs between cognitive load and motor complexity.

The tongue joystick has a direct spatial layout and thus the lowest cognitive load. While this spatial layout inherently biases our study in favor of the tongue joystick, it is the goal of the study to test whether the tongue has enough accuracy and agility to allow this lower cognitive load to dominate its performance.

The sip/puff switch, inversely, has the highest cognitive load. Its two-word prompts (e.g. "hard sip") must be combined along two dimensions (soft vs hard, puff vs sip) to produce a motor action.

The bite sensor also uses a combination prompt (e.g. "long short"), but each word maps to an independent motor action modulated only by time (short vs long). Thus the bite sensor should have a slightly lower cognitive load. The bite sensor also has the simplest motor action, a bite, requiring little physical agility.

We have the following hypotheses:

1. The joystick will be the fastest, as it has a direct mapping from prompt to motor action.
2. The bite sensor, with a smaller cognitive load, will start out faster than the sip/puff switch. After a learning period where the sip/puff prompts are internalized, the results should flip, as sip/puff signals are shorter.

Experimental Design

The sip/puff switch, bite sensor, and tongue joystick were compared using a within-subjects design, counterbalanced for device order. Twenty-four users (11 women and 13 men, ranging in age from 18 to 58) participated. Users were recruited from an on-campus behavioral research subject pool and were paid $15 for the one hour study.

Figure 6 and Figure 7 show the experimental setup. Devices were clamped at mouth height. Stimuli were presented on

Figure 6: A subject performs a forced-choice four-direction speed test of the tongue joystick.

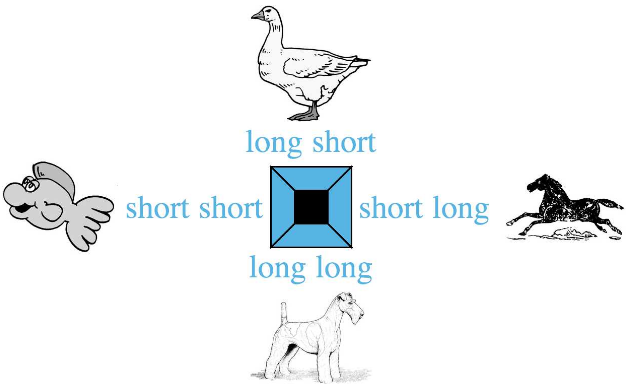

Figure 7: The Animals task with prompts for the bite sensor.

screen and consisted of four choices encircling a segmented central square with text reminders of the selection method. When a direction was chosen, feedback was provided by temporarily changing the color of the associated segment.

For each device in turn, subjects performed a Spelling and an Animal selection task. In the Spelling task, the device was explained and demonstrated to the subject, who then used it to spell six four-letter words. The correct letter occurred equally often in each of the four positions. Subjects were required to correct mistakes, thus ensuring a basic level of mastery of the device.

In the Animals task, we tested the response time of the device in combination with an audio prompt and changing selections, thus mimicking our target application. Subjects saw four new black and white animal pictures each trial (Figure 7) and were given an audio prompt, 250ms later, of "bird", "horse", "dog", or "fish". Although subjects were asked to do as many trials as possible in ten minutes, all subjects received fifteen blocks of four trials, with a four second rest period between blocks.

After performing both tasks, subjects filled out a device questionnaire with a seven point scale for each of speed (slow/fast), accuracy (inaccurate/accurate), lack of fatigue (tiring/not tiring), and ease of use (confusing/easy to understand).

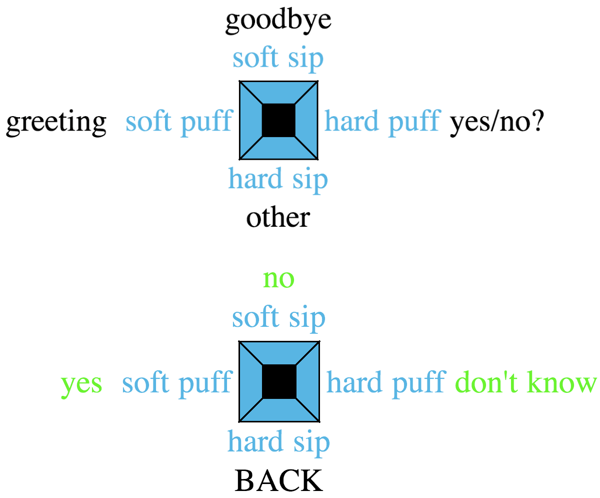

Figure 8: The two-level dialogue tree: category selection is shown at top; response selection for the category "yes/no?" is shown at bottom.

After testing each device, subjects performed a Dialogue task. Subjects were asked to choose one device to use for holding three conversations with a computer agent named Katie. Katie took four turns. For each of Katie's turns, the subject first assigned Katie's speech to one of the categories "greeting", "yes/no question", "goodbye", or "other" (Figure 8, *top*). Subjects then saw three within-category options for their response and the "back" option in case they wished to re-categorize Katie's speech (Figure 8, *bottom*). When one of the response options was chosen, an audio file with that meaning was played aloud, and the screen reset to the category level.

A final questionnaire asked for ranked comparisons of the three devices using the superlative of the four scales (fastest, most accurate, least tiring, and easiest to understand). Subjects also provided general comments about their experiences and preferences.

RESULTS

We give results of the Animals and Dialogue tasks, followed by summaries of the surveys.

Animals Task

The tongue joystick proved the fastest and most accurate in the Animals task. Means and standard deviations for reaction time, measured from the onset of the visual prompt, are reported in Table 1. A repeated measures ANOVA showed

Device	N	Reaction Time (ms)		Error Rate
		Mean	Std Dev	
Sip/puff	24	3186	1721	23%
Bite sensor	24	3016	1258	20%
Tongue joystick	24	2065	880	10%

Table 1: Reaction times for the Animals task for all subjects.

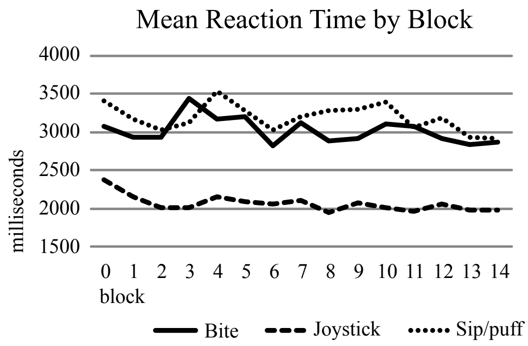

Figure 9: Average reaction times of the twenty-four subjects across the fifteen blocks of the Animals task.

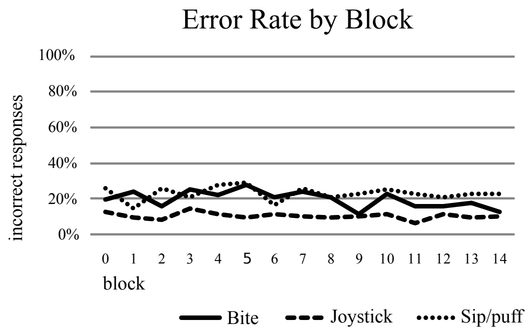

Figure 10: Average error rates of the twenty-four subjects across the fifteen blocks of the Animals task.

that device significantly affected reaction time ($F(2, 22) = 42.27$, $p < .001$). Bonferroni-corrected pairwise comparisons showed the joystick to be significantly faster than the bite sensor ($p < .001$) and the sip/puff switch ($p < .001$). No significant difference was found between the latter two devices. Thus only the first of our hypotheses, not the second, was borne out. We conjecture that subjects were still mastering the bite sensor and the sip/puff switch during the fifteen blocks; the variability in reaction times for these two devices, seen in Figure 9, supports this view. In contrast, subjects learned the tongue joystick quickly and had consistently less variability. In future work, we plan to extend our study, looking for trends over a longer time period.

The joystick was the most accurate (Bonferroni-corrected, $p = .024$ vs bite, $p = .001$ vs sip/puff); no significant difference in error rates was found between the other two devices. Error rates by block are given in Figure 10. An analysis of the confusion matrices revealed that subjects had the most trouble with the "short long" and "long short" prompts for the bite sensor, and with thresholding soft versus hard for the sip/puff switch. This result supports our conjecture that subjects were still in transition for the two devices.

Dialogue Task

The Dialogue task recreates a language understanding component closer to our target application. For this task, 17 (71%) of the subjects chose to use the joystick, 5 (21%) chose the sip/puff switch, and 2 (8%) chose the bite sensor.

	Category (action)	Response Latency (ms)		Error
		Mean	Std Dev	
Social	greeting (left)	3552	1832	0%
	bye (up)	3412	1582	6%
Content	yes/no (right)	4129	2695	14%
	other (down)	4198	2494	20%

Table 2: Average milliseconds taken by the twelve joystick subjects to navigate through the two-level tree, grouped by top level (category) and type of response (social vs content).

Error rates were computed based only on category selection; Katie would reply cooperatively to any second-level response. Several of the subjects simply did not understand the task; because our interest is the viability of the device when used by a skilled performer, we eliminated from our analysis any subject whose error rate was above 25%. This elimination left 12 joystick users, 1 sip/puff user, and 2 bite sensor users. With such small numbers for two devices, we analyze only the tongue joystick results.

Table 2 shows mean response latencies and error rates for the tongue joystick, grouped by category. Latency is defined as the time from the end of Katie's speech, through category selection, to final response selection.

The difference in response latency between social and content categories was not significant. Error rates were significantly different between social and content categories ($F(1, 11) = 12.78$, $p = .004$). This significance is not attributable to direction, which showed no comparable pattern in the Animals task.

Overall, subjects took about 3.8 seconds to compose a response (two selections at about 2 seconds per selection), with 90% accuracy, roughly in line with their Animals task results.

User Surveys

Figure 11 shows the mean ratings for the individual surveys. On every dimension, the joystick was rated significantly better than the bite sensor (Bonferroni-corrected, $p < .05$ for each) and comparable to or significantly better than the sip/puff switch ($p < .05$ for accuracy only). Users did not rate the joystick significantly faster than the sip/puff switch, despite the better joystick timings.

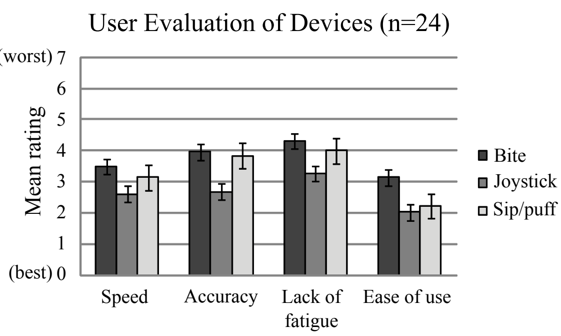

Figure 11: Average ratings given by subjects to each device after testing it in the Animals task.

Ordering	Fastest	Most accurate	Least tiring	Easiest
joy > bite > sip	8	10	4	9
joy > sip > bite	2	4	2	4
bite > joy > sip	2		2	1
bite > sip > joy		1		
sip > joy > bite	3	2	7	
sip > bite > joy	3	1	3	4

Table 3: Frequencies of the six device orderings for each category in the final comparative rankings (18 subjects completed the survey correctly).

After the Dialogue task, users provided a set of comparative rankings, tabulated in Table 3. Preferences at the end of the study mirror earlier judgments. The joystick was ranked better than the bite sensor for all characteristics (Wilcoxon Signed Rank, $p < .05$ for all), but was ranked significantly higher than the sip/puff device only for accuracy ($p < .01$). No comparisons between the bite sensor and sip/puff switch were significant.

ARTICULATED CHARACTER HEAD PROTOTYPE SYSTEM

We incorporated the tongue joystick into a system prototype using the custom articulated-head dragon shown in Figure 1. We built a heads-up display that lodges in the character's snout, and designed a dialogue tree appropriate for interacting with a young child (Figure 12). The actor holds the tongue joystick in her mouth while donning the dragon head. When she presses one of the petals in the tongue joystick, feedback is shown on the heads-up display; if a line of audio is triggered, it plays out of a speaker while the mouth moves automatically to prerecorded puppeteering. All hardware is connected through an external laptop.

As shown in the accompanying video, where the dragon meets a 6-year-old girl, the actor is able to use the tongue joystick quickly enough to engage in smooth, natural turntaking using prerecorded audio. The dialogue-tree system excels at constrained situations where responses are predictable, such as knock-knock jokes and simple questions such as "what's your favorite color?". We thus optimized the tree for taking the initiative to control the conversation and guide it down one of these paths. This situation mirrors that of our target application, where conversations with characters in the theme parks are usually character driven. The tree also had a branch containing responses such as yes, no, and thanks.

From our trial sessions we learned that we need to more thoroughly examine the dynamic of conversation with children. We rapidly learned that one often needs to repeat a phrase until they focus enough to hear. Also, having a supply of giggles and other phatic expressions that could be inserted into a dialogue path without interrupting it would cover most awkwardnesses.

DISCUSSION AND CONCLUSION

We presented a new tongue input device, the tongue joystick, for maneuvering within a dialogue tree to select pieces of prerecorded audio. The device has a pacifier form factor which makes it both user-independent and firmly grippable for speed and accuracy. In building the tongue joystick, we developed a new method of constructing rugged, soft, low-actuation force devices by using both molding and sewing techniques to combine soft silicone and flexible conductive thread. Our hope is that this method will be useful to the assistive technology community in designing new devices.

The form factor and manufacturing method of our tongue joystick are amenable to easily incorporating other sensors: a bite switch could be located in the stem, for example, or a sip/puff tube embedded in the middle. Airflow and weight must be considered in any design, however. The device fit inside the dragon head, but the wires reduced the head mobility of the actor; we plan to build our next prototype with a wireless chip and small battery embedded in the cheek guard.

Our study design attempted to balance speed and accuracy. We urged subjects to work quickly, but did not penalize them for wrong answers. The error rate for the tongue joystick, 10%, was thus very high. In future work, we will run studies that isolate reaction time and error rate, giving a better lower bound for each while helping us better understand learning and fatigue effects. Future studies will also train users longer so that they will be closer to the "expert" user expected in our application.

The current user study showed that subjects were able to use the tongue joystick to respond to conversational turns with pauses around 3-4 seconds. Our articulated character head prototype system showed that a skilled actor could control conversational turns even faster. The maximum amount of latency that avoids conversational awkwardness is situation dependent; for instance, while cross-cultural analysis has shown that adults minimize the silence between turns [28], at least one study has demonstrated that 7-10 year olds will converse cooperatively with an animated character with a 3-4 second response latency. Our application has more resilience than standard conversation, as the actor can cover pauses with physical acting. In our prototype testing, a case arose where the actor did not have audio for a particular situation; she switched from dialogue to solely physical acting for several minutes without the child noticing the transitions. Thus with a combination of lively acting, expressive costumes, and our tongue joystick, we hope to bring a new set of characters to life.

REFERENCES

1. L. Andreasen Struijk. An inductive tongue computer interface for control of computers and assistive devices. *IEEE Transactions on Biomedical Engineering*, 53(12):2594–2597, 2006.

2. M. Betke, J. Gips, and P. Fleming. The camera mouse: visual tracking of body features to provide computer access for people with severe disabilities. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 10:1–10, 2002.

3. J. Bilmes, X. Li, J. Malkin, K. Kilanski, R. Wright, K. Kirchhoff, A. Subramanya, S. Harada, J. Landay, P. Dowden, and H. Chizeck. The vocal joystick: a voice-based human-computer interface for individuals with motor impairments. In *Conference on Human*

```
3 (chcknjoke) "I have a joke. Do you want to hear it?"
  1 (Yes:Xroad?) "Why did the dragon cross the road?"
    1 (punch) "Because...chickens weren't invented yet!"
    2 (answr:hey) "Hey, how did you know that?"
    3 (nope:punch) "Nope! 'Cause chickens weren't invented yet!"
  2 (No:Sad) "No? But it's such a funny joke. I'm sad."
```

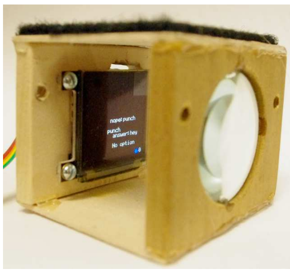

Figure 12: A snippet of dialogue tree, shown on the heads-up display normally lodged in the dragon's snout. The display contains a collimating lens so that the actor's eyes do not need to refocus between the screen and child.

Language Technology and Empirical Methods in Natural Language Processing, pages 995–1002. Association for Computational Linguistics, 2005.

4. Broadened Horizons. Bite Switch, 2011. `broadenedhorizons.com`.

5. Broadened Horizons. QuadMouse, 2011. `broadenedhorizons.com`.

6. C. Darves and S. Oviatt. Adaptation of users' spoken dialogue patterns in a conversational interface. In *International Conference on Spoken Language Processing*, pages 561–564, 2002.

7. P. DiMattia and J. Gips. Eagleeyes: Technologies for non-verbal persons. In K. Thies and J. Travers, editors, *Handbook of Human Development for Health Care Professionals*, pages 429–448. Jones and Bartlett, 2005.

8. EnableMart. EyeTech TM2, 2009. `enablemart.com`.

9. G. Evreinov and T. Evreinova. "Breath-joystick" – graphical manipulator for physically disabled users. In *International Conference on Computers Helping People with Special Needs*, pages 193–200, 2000.

10. T. Felzer and B. Freisleben. Hawcos: the "hands-free" wheelchair control system. In *Conference on Assistive Technologies*, pages 127–134. ACM, 2002.

11. L. Hochberg and J. Donoghue. Sensors for brain-computer interfaces. *IEEE Engineering in Medicine and Biology Magazine*, 25(5):32–38, Sept.-Oct. 2006.

12. X. Huo and M. Ghovanloo. Using constrained tongue motion as an alternative control surface for wheeled mobility. *IEEE Transactions on Biomedical Engineering*, 56(6):1719–1726, 2009.

13. T. Igarashi and J. Hughes. Voice as sound: using non-verbal voice input for interactive control. In *Symposium on User Interface Software and Technology*, pages 155–156. ACM, 2001.

14. J. S. Kim, H. Jeong, and W. Son. A new means of HCI: EMG-MOUSE. In *IEEE Conference on Systems, Man and Cybernetics*, volume 1, pages 100–104, 2004.

15. K. Koichi. Input device for disabled persons using expiration and tooth-touch sound signals. In *Symposium on Applied Computing*, pages 1159–1164. ACM, 2010.

16. T. Kuiken, G. Li, B. Lock, R. Lipschutz, L. Miller, K. Stubblefield, and K. Englehart. Targeted muscle reinnervation for real-time myoelectric control of multifunction artificial arms. *Journal of the American Medical Association*, 301(6):619–628, 2009.

17. J. LaCourse and F. Hludik. An eye movement communication-control system for the disabled. *IEEE Transactions on Biomedical Engineering*, 37(12):1215 –1220, 1990.

18. F. Loewenich and F. Maire. Hands-free mouse-pointer manipulation using motion-tracking and speech recognition. In *Australasian Conference on Computer-Human Interaction: Entertaining User Interfaces*, pages 295–302. ACM, 2007.

19. F. Masuda and C. Wada. Effects of visual stimuli on a communication assistive method using sympathetic skin response. In *International Conference on Computers Helping People with Special Needs*, pages 189–192. Springer-Verlag, 2010.

20. M. Moore and U. Dua. A galvanic skin response interface for people with severe motor disabilities. In *Conference on Computers and Accessibility*, pages 48–54. ACM, 2004.

21. Natural Point. SmartNav, 2011. `naturalpoint.com`.

22. New Abilities. Toungue-touch keypad, 2009. `newabilities.com`.

23. Origin Instruments Corporation. Sip/puff Breeze, 2011. `orin.com`.

24. S. Patel and G. Abowd. BLUI: low-cost localized blowable user interfaces. In *Symposium on User Interface Software and Technology*, pages 217–220. ACM, 2007.

25. C. Salem and S. Zhai. An isometric tongue pointing device. In *Conference on Human Factors in Computing Systems*, pages 538–539. ACM, 1997.

26. T. S. Saponas, D. Kelly, B. Parviz, and D. Tan. Optically sensing tongue gestures for computer input. In *Symposium on User Interface Software and Technology*, pages 177–180. ACM, 2009.

27. A. Sears, M. Young, and J. Feng. Physical disabilities and computing technologies: An analysis of impairments. In A. Sears and J. Jacko, editors, *The Human-computer interaction handbook*, pages 829–852. 2008.

28. T. Stivers, N. Enfield, P. Brown, C. Englert, M. Hayashi, T. Heinemann, G. Hoymann, F. Rossano, J. Peter de Ruiter, K. E. Yoon, and S. Levinson. Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, 106(26):10587–10592, 2009.

29. J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, and T. Vaughan. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6):767–791, 2002.