

Article

Sensing the Neck: Classifying Movements and Actions with a Neck-Mounted Wearable Device [†]

Jonathan Lacanlale, Paruyr Isayan, Katya Mkrтчyan  and Ani Nahapetian *

Computer Science Department, California State University, Northridge (CSUN), Northridge, CA 91330, USA; jonathan.lacanlale.608@my.csun.edu (J.L.); paruyr.isayan.536@my.csun.edu (P.I.); katya.mkrтчyan@csun.edu (K.M.)

* Correspondence: ani@csun.edu

[†] This is an extended version of conference paper Lacanlale, J.; Isayan, P.; Mkrтчyan, K.; Nahapetian, A. Look Ma, No Hands: A Wearable Neck-Mounted Interface. In Proceedings of the Conference on Information Technology for Social Good, Roma, Italy, 9–11 September 2021.

Abstract: Sensor technology that captures information from a user's neck region can enable a range of new possibilities, including less intrusive mobile software interfaces. In this work, we investigate the feasibility of using a single inexpensive flex sensor mounted at the neck to capture information about head gestures, about mouth movements, and about the presence of audible speech. Different sensor sizes and various sensor positions on the neck are experimentally evaluated. With data collected from experiments performed on the finalized prototype, a classification accuracy of 91% in differentiating common head gestures, a classification accuracy of 63% in differentiating mouth movements, and a classification accuracy of 83% in speech detection are achieved.

Keywords: wearable computing; interaction design; neck-mounted interface; flex sensor; machine learning (ML)



Citation: Lacanlale, J.; Isayan, P.; Mkrтчyan, K.; Nahapetian, A. Sensing the Neck: Classifying Movements and Actions with a Neck-Mounted Wearable Device. *Sensors* **2022**, *22*, 4313. <https://doi.org/10.3390/s22124313>

Academic Editors: Claudio Palazzi, Ombretta Gaggi, Pietro Manzoni and Mehmet Rasit Yuce

Received: 6 March 2022

Accepted: 30 May 2022

Published: 7 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The ever-increasing prevalence of mobile phones, wearable devices, and smart speakers has spurred intense exploration into user interfaces. These new user interfaces need to address the challenges posed by the ubiquitous interaction paradigm, while having available the possibilities that these varied smart technologies provide.

Arenas for exploration of mobile user interfaces include improving gesture-based interfaces to enable interaction in limit mobility settings or by decreasing the social disruption that is caused by repeated disruptive interactions. Interfaces have been developed that use the movement of the hands, arms, eyes, and feet.

Touch gesture controls still dominate mobile system interfaces because of the ubiquity of touch screens [1]. However, the dominant tap, scroll, and pinch gestures have been linked to repetitive strain injuries on smart phones [2,3]. In addition, they have their limitations on wearable devices because of the limited screen size and, in turn, the available interface surface. The gestures on smartwatch screens need to be done with greater precision and with more constriction of the hand muscles, since the smartwatch screens are significantly smaller than the smartphone screens.

Voice user interfaces (VUIs) that are used for smart speakers have been another arena for improvement, with voiceless speech being explored for situations where there is background noise and for microinteractions.

In this work, we examine the benefits that sensing the neck can provide within the breadth of mobile user interfaces. We explore and develop a new user interface for mobile systems, independent of limb motions. For example, in place of a scroll down, the head can be tilted forward. In place of a tap, the head can be turned to one side, all with only an inexpensive sensor affixed to the neck or shirt collar.

We sensor the neck with an inexpensive and nonintrusive flex sensor and show the range of interfaces that are possible with the incorporation of this simple wearable technology into our lives. Our efforts provide a proof of concept that common actions, such as head tilts, mouth movements, and even speech, can be classified through the interpretation of the bend angle received from the neck. We explore the size of the flex sensor and the positioning of the sensor on the neck and use our classification results to tailor the prototype.

Applications for neck interfaces include use in assistive devices where limb motion is limited, in gaming and augmented reality systems for more immersive experiences, and in wearable and vehicular systems where hand and/or voice use is restricted or inconvenient. Neck interactions expand a user's bandwidth for information transference, in conjunction with or in place of the typically saturated visual and the audial channels.

A neck-mounted prototype was designed and developed, as detailed in Section 3. The system design considered comfort and the range of motion in the neck and upper body. The form factor and the positioning of the system was finalized to enable the embedding in clothing, such as in a shirt collar. A range of sensor types, sizes, and positions were considered and evaluated.

The prototype's head gesture and position classification accuracy was evaluated for five different classes of common head tilt positions. These experimental evaluations are detailed in Section 4. Head tilt classification is important because it enables user interface input with simple and subtle head gestures.

The encouraging results from the head gesture classification motivated us to explore more possibilities, including using the prototype for mouth movement and speech classification. The experimental evaluations of mouth movements and speech classification are detailed in Section 5. By also incorporating speech and/or mouth movement detection, head gestures for software interactions can be differentiated from head gestures that arise during regular conversation.

The main contributions of this work are (1) the development of a neck-mounted prototype, with an evaluation of sensor types, sizes, and positions; (2) the evaluation of the prototype's head-position classification accuracy; (3) mouth movement detection; and (4) speech detection and classification.

2. Related Work

Interfaces that sense hand and arm gestures are widespread [4], including those that rely on motion sensors [5–8], changes in Bluetooth received signal strength [9], and light sensors [10,11]. Interfaces that leverage the movement of the legs and the feet have also been explored [12,13]. Computer vision-based approaches using the camera to capture head and body motions [14,15], facial expressions [16], and eye movement [17] also exist.

Detection of throat activity has been explored using different enabling technologies. Acoustic sensors have been used for muscle movement recognition [18], speech recognition, ref. [19] and actions related to eating [20–22]. Prior research has been done on e-textiles used in the neck region for detecting posture [23] and swallowing [24], but those efforts have relied on capacitive methods that have limitations in daily interactions. Researchers have explored sensing the neck with piezoelectric sensors for monitoring eating [25] and medication adherence [26].

In addition to the neck-mounted sensors systems, there has been an exploration of actuation at the neck region using vibrotactile stimulation for accomplishing haptic perception [27–29].

The use of video image processing for speech recognition has been applied to lip reading [30–32]. More recently, as part of the silent or unvoiced speech recognition research efforts, mobile phone and wearable cameras have been used for speech classification from mouth movements. Researchers have used bespoke wearable hardware for detecting mouth and chin movements [33], or leveraged smart phone cameras [34].

Electromyography (EMG) has also been used for speech and/or silent speech classification. Researchers have used EMG sensors on the fingers placed on the face for mouth movement classifications [35]. EMG sensing of the face for speech detection has also been carried out [36].

Tongue movement has been monitored for human–computer interfaces, including using a magnetometer to track a magnet in the mouth [37], using capacitive touch sensors mounted on a retainer in the mouth [38], using EMG from the face muscles around the mouth [39], and using EMG coupled with electroencephalography (EEG) as sensed from behind the ear [40]. Detecting tooth clicks has also been explored including a teeth-based interface that senses tooth clicks using microphones placed behind the ears [41].

Head position classification has been carried out with motion sensors on the head [42], pairing ultrasound transmitters and ultrasonic sensors mounted on the body [43] and barometric pressure sensing inside the ear [44].

This work is an expansion on our previously published conference paper [45] that classified head gestures using on a single neck-mounted bend sensor. In this expanded work, we look not only at head gesture classification using our neck-mounted sensor interface, but also at mouth movement classification, speech detection, and speech classification.

3. Prototype

A neck-mounted wearable prototype was developed and used for classifying neck movement, mouth movement, and speech. The prototype consists of a sensor affixed to the neck which is connected to a microcontroller. The data collected from the sensor is wirelessly transferred via Bluetooth by the microcontroller to the user’s paired smart phone. On the smart phone, the time-series data is in real time filtered, classified, and then used as input to a software application. Figure 1 provides an overview of the wearable system and its components interactions.

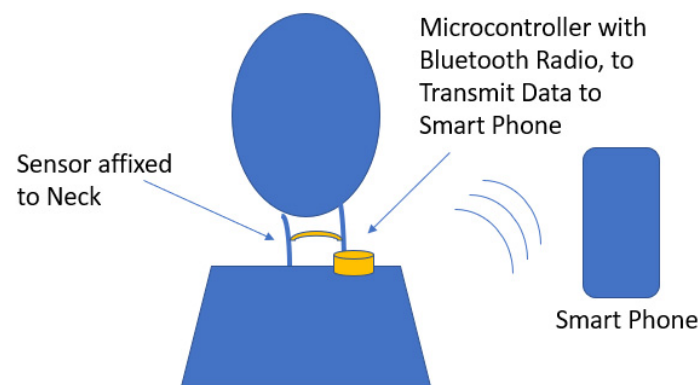


Figure 1. Prototype system’s component overview, with sensor placed on neck and wearable hardware placed on collar for communicating data to a smartphone for processing and for interfacing with the application.

E-textile and flex sensors were investigated as potential candidates for the prototype. E-textiles can be used as capacitive sensors or as resistive sensors. With the capacitive method, the e-textile worked well as a proximity sensor to detect when the sensor was near human skin. However, once the sensor was in contact with or in close proximity of the skin, the sensor data became saturated and did not provide valuable features or respond to movements. Using the e-textile sensor as a resistive sensor was more successful in displaying features when actively bending or pulling the material.

The flex sensor proved to be the most appropriate for sensing the neck. The flex sensor acts as a flexible potentiometer, whose resistance increases as the bend angle increases. Unlike the e-textile, which did not return to a static level after deformation and was prone to noise, the flex sensor performed reliably under bending and returned to a stable level when straight.

A variety of positions for the sensor around the neck, chin, and side of face were explored with the neck being the most practical in terms of data collection and ease of wear.

The hardware of the final prototype consists of an inexpensive (approximately USD 10) flex sensor, whose change in resistance signaled change in the bend of the sensor. The flex sensor was placed against the neck by weaving it under a small piece of paper that was taped to the neck. An Arduino microcontroller collected and wirelessly transmitted the data from the sensor to a smart phone for processing and display. Both an Arduino Nano and an Arduino Mega 2560 were used in the experiments.

A simple moving average (SMA) filter was used to smooth the measured resistance signal. SMA filters replace the current data value with the unweighted mean of the k previous points in the data stream, in effect smoothing the data by flattening the impact of noise and artifact that is outside the bigger trend of the data. As the window size is decreased, the smoothness of the data is decreased. In this application, a window size that is too small can result in artifact and/or noise in the time-series data being improperly classified as a neck movement event. As the window size is increased, the impact of noise and artifact is also decreased, but the likelihood that relevant information is filtered out is increased. In this application, with a window size that is too large, there is the risk of delaying the recognition of neck movement events or even missing the events altogether. A window size of $k = 40$ was selected, which roughly maps to one second of data.

4. Head Tilt Detection

In a series of experiments, two types of flex sensors in a variety of positions on the neck are evaluated to determine the feasibility of differentiating and classifying head tilt and positioning.

In the experiments conducted, both a short sensor in three different positions and a long sensor were considered. Each sensor placement and sensor received 10 experiments per head-tilt with a time duration of 30 s. The tilts were held static for the entire 30 s. For each experiment, approximately 1100 data points were collected.

4.1. Flex Sensor Types and Placement

Two types of flex sensors are considered: a short sensor and a long sensor. With the short sensor, three different placements are considered: a low placement, a center placement, and a high placement. The low placement is at the bottom of the neck, closest to the collar, as shown in Figure 2a. The center placement is directly over the larynx, at the middle of the neck, as shown in Figure 2b. The high placement is the top of the throat, closest to the chin, as shown in Figure 2c. The long sensor spans the three positions along the neck, from the base of the neck to under the chin, as shown in Figure 3.

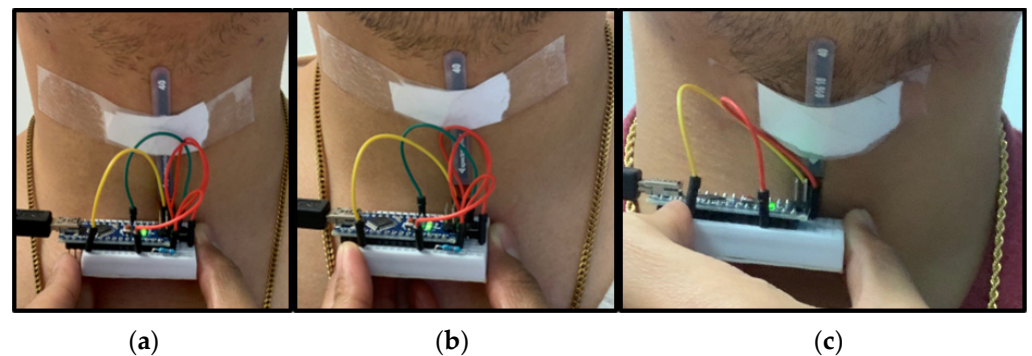


Figure 2. (a) Low, (b) center, and (c) high placement of the short flex sensor along the center line of the neck.

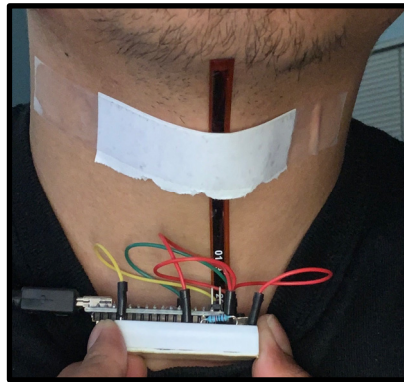


Figure 3. The placement of the long flex sensor along the center line of the neck.

4.2. Data Visualization

We visualize here some of the data collected across various placements of the sensors and for different head tilts. Figures 4–6, respectively, display the collected resistance data over a 30-s time frame across the first three classes of head tilts, namely down, forward/no tilt, and up, for each placement of the short sensor, namely low, center, and high placement. Figure 7 displays the collected resistance data over a 30-s time frame for the long sensor, across the first three classes of head tilts, namely down, forward, and up. The data represented has been filtered using a moving average filter.

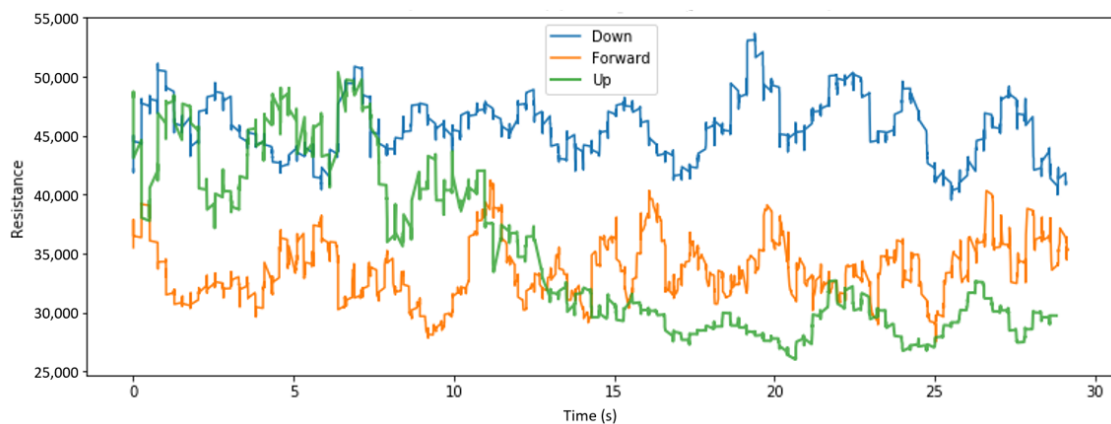


Figure 4. With low placement of short sensor, head tilt filtered data.

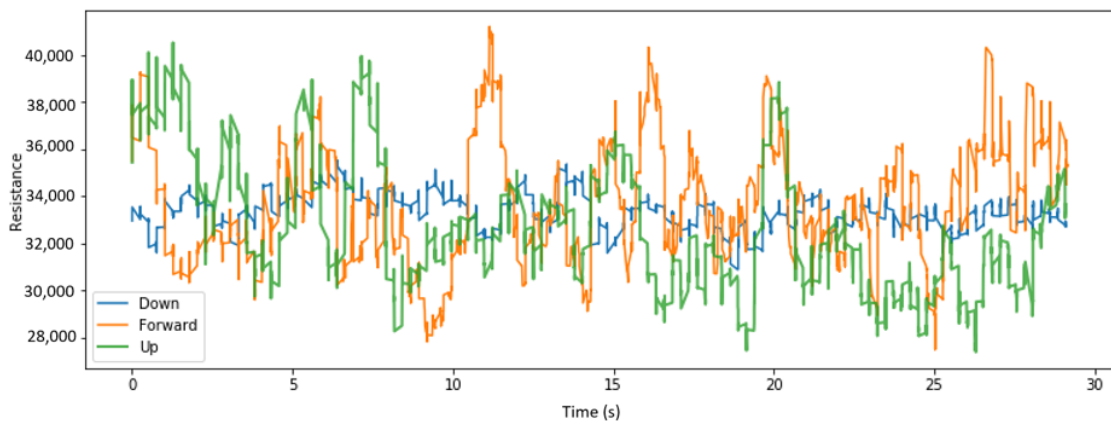


Figure 5. With center placement of short sensor, head tilt filtered data.

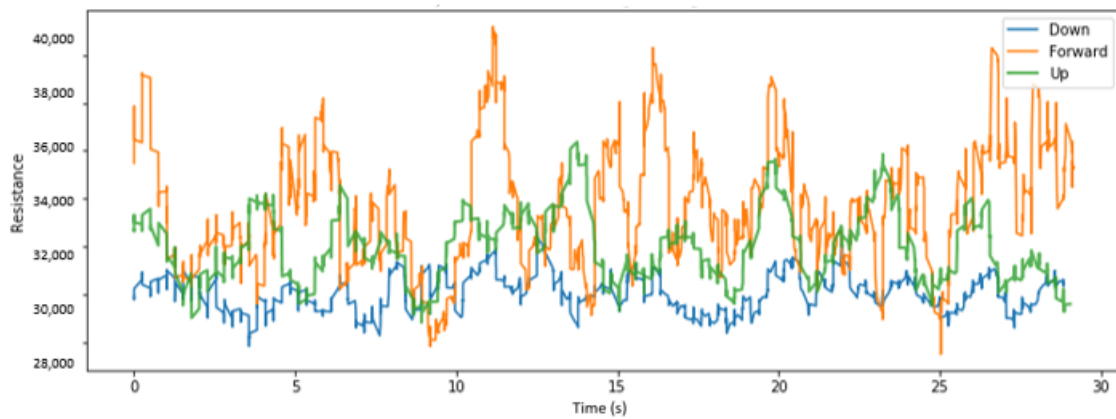


Figure 6. With high placement of short sensor, head tilt filtered data.

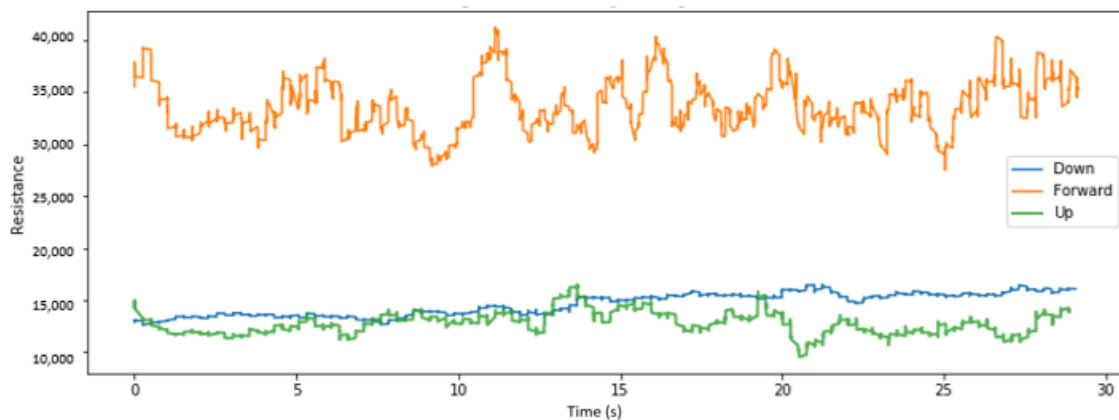


Figure 7. With long sensor, head tilt filtered data.

The short, low sensor placement and the long sensor (Figures 4 and 7, respectively) show the clearest distinction between the three classes. Therefore, the short, low sensor placement and the long sensor were further evaluated using all five classes of head tilts, namely down, forward, up, right, left. The collected resistance data over a 30-s time frame are shown in Figures 8 and 9, respectively.

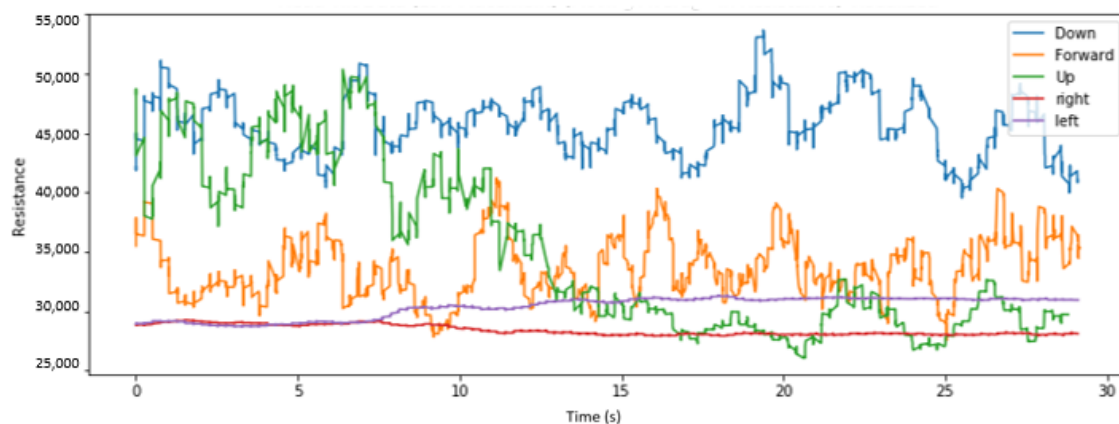


Figure 8. With low placement of short sensor, head tilt filtered data, with right and left tilts added.

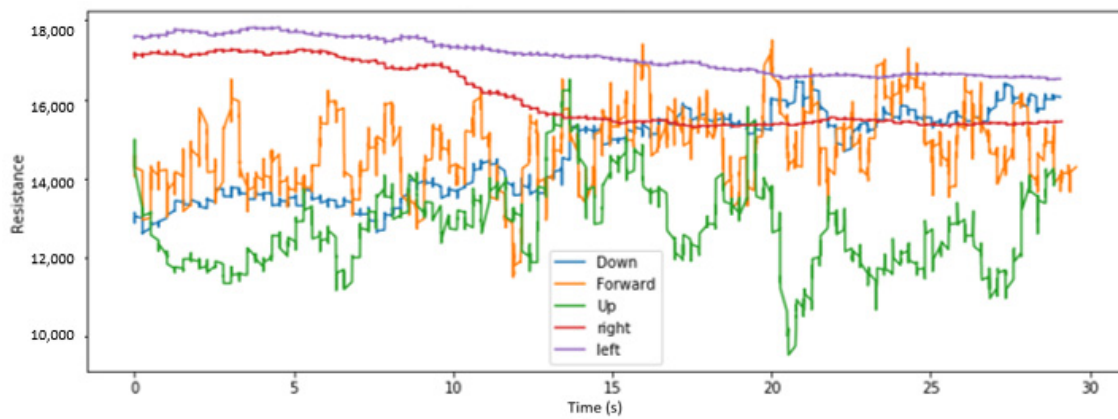


Figure 9. With long sensor, head tilt filtered data, with right and left tilts added.

4.3. Head Tilt Detection Machine Learning Results

We evaluated the accuracy of classifying a three-class dictionary of head tilts. We then went on to evaluate the accuracy of classifying an expanded five-class dictionary of head tilts. The classification results are presented in this subsection.

Three different classical machine learning (ML) classifiers were considered, specifically logistic regression, SVM, and random forest. The labeled dataset was partitioned into a train and held-out test set with an 80:20 ratio. To ensure the consistency of the models, a k -fold cross-validation was performed. A fivefold cross-validation of the train set was performed, with a random fourth of the examples in the training fold being used for validation during hyper-parameter tuning. For all the classical ML models, the Scikit-learn library in Python was used.

All four configurations, i.e., the long sensor and the three (low, center, and high) placements of the short sensor, were evaluated using the three head tilts (down, forward/not tilt, and up).

Table 1 displays our fivefold accuracy based on the model and placements of the sensors. In all cases, Logistic Regression was not sufficient in classifying the three-class dictionary. The short and low sensor placement and the long sensor had the best results. In both cases, random forest is the best performing model with test accuracies reaching ~83.4% and ~96% for the short, low placement and the long sensor, respectively.

Table 1. Fivefold training, cross-validation, and held-out test accuracy of classical ML models with different feature sets. The bold font denotes the cases with the highest accuracy for that model. These results are for the three-class dictionary.

Model		Short Sensor Low Placement	Short Sensor Center Placement	Short Sensor High Placement	Long Sensor
Logistic Regression	Train	0.744	0.379	0.629	0.603
	Validate	0.74	0.379	0.622	0.602
	Test	0.76	0.349	0.589	0.608
SVM	Train	0.825	0.594	0.648	0.891
	Validate	0.809	0.547	0.612	0.881
	Test	0.824	0.555	0.575	0.891
Random Forest	Train	0.955	0.918	0.854	0.989
	Validate	0.821	0.665	0.694	0.945
	Test	0.834	0.669	0.671	0.960

To the best performing results, two additional classes were added. The two additional classes are the user's head facing right and the user's head facing left.

Table 2 shows the performance of the short sensor with low placement and the long sensor when classifying against this five-class dictionary. As with previous results, random forest had the best performance with a test accuracy of ~83% for the short sensor and ~91% for the long sensor.

Table 2. Fivefold training, cross-validation, and held-out test accuracy of classical ML models with different feature sets. The bold font denotes the cases with the highest accuracy for that model. These results are for the five-class dictionary that includes facing right and facing left.

Model		Short Sensor Low Placement	Long Sensor
Logistic Regression	Train	0.734	0.337
	Validate	0.733	0.338
	Test	0.755	0.363
SVM	Train	0.756	0.869
	Validate	0.741	0.812
	Test	0.76	0.818
Random Forest	Train	0.956	0.977
	Validate	0.824	0.915
	Test	0.828	0.91

Table 3 shows the confusion matrix for the short sensor with low placement with the random forest classifier. The largest source of misclassifications are from the up data points, with only 65 out of 157 labels predicted correctly.

Table 3. Five-class confusion matrix for the short sensor with low placement. Rows represent actual class and columns represent predicted class.

Random Forest		Predicated				
		Down	Forward	Up	Right	Left
Actual	Down	259	0	10	0	0
	Forward	1	285	40	1	3
	Up	25	47	65	15	5
	Right	0	0	8	185	18
	Left	0	0	3	29	194

Table 4 shows the confusion matrix for the long sensor using the random forest classifier. With the long sensor, only 17 out of 182 up data points are mislabeled. The largest confusion is between left and right tilts.

From the confusion matrix the neck gesture language can be created. The most frequent or the most important gestures can be assigned to the head tilts that achieve the highest classification accuracy, both in terms of sensitivity and specificity. For example, the following mapping of neck gestures would be appropriate for the social media app Instagram. While on their feeds, users would tilt their heads forward to signal scrolling and would turn their heads to the side, either right or left, to 'like' an image.

Table 4. Five-class confusion matrix for the long sensor. Rows represent actual class and columns represent predicted class.

Random Forest		Predicated				
		Down	Forward	Up	Right	Left
Actual	Down	202	3	11	0	0
	Forward	0	494	2	0	0
	Up	17	0	182	0	0
	Right	0	0	0	204	49
	Left	0	0	0	36	166

5. Speech and Mouth Movement Detection

In this section, we explore a larger range of opportunities that the neck-mounted sensor can provide in addition to the head gesture detection detailed in Section 4. Section 5.1 addresses speech detection using the prototype, by differentiating speech from static breathing. Section 5.2 address mouth movement classification, namely the determination of how many times the mouth has been opened and closed. Section 5.3 tackles the challenging task of speech classification using only the detection of movement in the neck.

Speech and mouth movement detection provide contextual information that can be used to trigger or to mute the head tilt interface. For instance, if the system detects that the user is talking, then the user's head tilts are not relayed to application software.

5.1. Speech Detection

Figure 10 shows an example sensor reading from static breathing and from talking, specifically saying 'hello', on the same graph. The visualization demonstrates that the presence of speech can potentially be differentiated from static breathing using only the data collected from the flex sensor on the neck-mounted prototype.

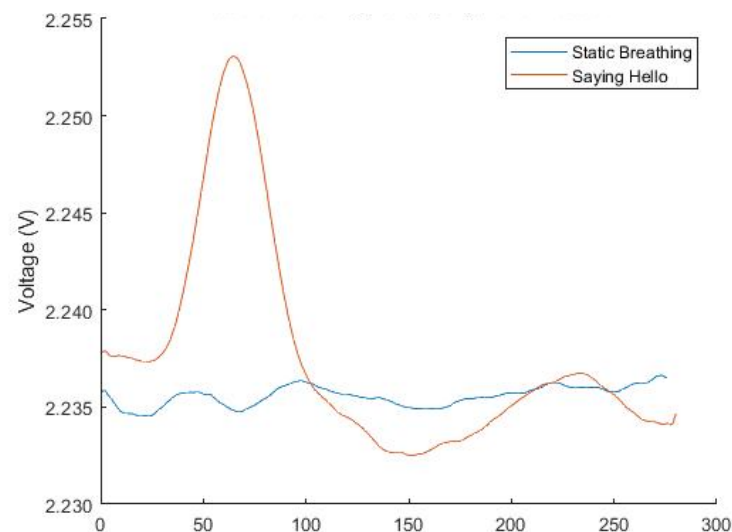


Figure 10. Sensor readings from static breathing and saying 'hello'.

Using the neck-mounted prototype, an experiment was conducted to see if static breathing can indeed be differentiated from speech. Three-second-long samples with the prototype's flex sensor were collected of both static breathing and of saying 'hello'. A total of 60 samples, 30 of each class, were collected. The samples were classified using K-nearest neighbors (k-NN) with dynamic time warping (DTW), with k set to 3.

Dynamic time warping measures the similarity between two time-series signals, which may vary in speed and in length. It calculates the minimal distance between the signals

allowing for warping of the time axis, with similar signals having lower cost than dissimilar signals.

Each test signal is compared against all the training signals, and the DTW cost between the test signal and each training signals is calculated. The DTW cost of the k nearest neighbors, i.e., most similar training signals, is then used to classify the signal.

Table 5 shows the confusion matrix for the classification results. The overall accuracy of the classification was 83.3% with 3 of the 30 talking samples misclassified as breathing.

Table 5. Two-class confusion matrix for static breathing and talking. Rows represent actual class and columns represent predicted class.

		Predicated	
		Static Breathing	Talking
Actual	Static Breathing	23	7
	Talking	3	27

5.2. Mouth Movement Classification

In another experiment, the classification of mouth movements without the generation of any sound was examined. The mouth was opened and closed without sound being generated. It was a four-class dictionary, with static breathing (no mouth movement), opening and closing of the mouth once, opening and closing of the mouth twice, and opening and closing of the mouth three times.

Three-second-long samples with the prototype's flex sensor were collected with a total of 60 samples, 15 of each class. The samples were classified using K-nearest neighbors (k-NN) with dynamic time warping, with k set to 3.

Table 6 shows the confusion matrix for the classification results. The overall accuracy of the classification was 67.5%. The classification of static breathing resulted in most of the misclassifications. By considering sample's peak-to-valley amplitude, this misclassification can be decreased.

Table 6. Four-class confusion matrix for mouth movements. Rows represent actual class and columns represent predicted class.

		Predicated			
		Breathing	One Cycle	Two Cycles	Three Cycles
Actual	Breathing	2	3	3	12
	One cycle	0	19	1	0
	Two cycles	0	7	13	0
	Three cycles	0	0	0	20

5.3. Speech Classification

The final experiments explored speech classification. Two different experiments of speech classification were carried with each having a set of four different sentences or phrases being spoken with the prototype affixed to the neck and the bend sensor capturing the neck activity.

For each of the two experiments, three-second-long samples with the prototype's flex sensor were collected. For the first experiment with sentences, a total of 40 samples were collected, 10 of each class. The sentences used in the experiments were "I am a user who is talking right now"; "This is me talking with a sensor attached"; "Who am I talking to at this very moment?"; and "Can you recognize what I am saying while attached to a sensor?" For the second experiment with famous idioms, a total of 80 samples were collected, 20 of each class. The idioms used in the experiment were "a blessing in disguise"; "cut somebody some slack"; "better late than never"; and "a dime a dozen." The samples were classified using K-nearest neighbors (k-NN) with dynamic time warping, with k set to 3.

Tables 7 and 8 show the confusion matrices for the classification results for the two experiments, respectively. The overall accuracy of the classification was 62.5% and 32.5%, respectively.

Table 7. Four-class confusion matrix for spoken sentences. Rows represent actual class and columns represent predicted class.

		Predicated			
		"I Am ... "	"This Is ... "	"Who ... "	"Can You ... "
Actual	"I am a user who is talking right now."	0	9	1	0
	"This is me talking with a sensor attached."	0	10	0	0
	"Who am I talking to at this very moment?"	0	4	6	0
	"Can you recognize what I am saying while attached to a sensor?"	0	0	1	9

Table 8. Four-class confusion matrix for spoken phrases. Rows represent actual class and columns represent predicted class.

		Predicated			
		"A Blessing in Disguise"	"Cut Somebody Some Slack"	"Better Late than Never"	"A Dime a Dozen"
Actual	"A blessing in disguise"	0	0	14	6
	"Cut somebody some slack"	0	2	1	17
	"Better late than never"	0	0	19	1
	"A dime a dozen"	0	0	15	5

6. Discussion

The experiments with sensor data captured from the neck-mounted prototype show that the short sensor with low placement on the neck and the long sensor had the best results. For a three-class dictionary of head tilts, random forest is the best performing model with test accuracy of ~83.4% for the short sensor with low placement and ~96% for the long sensor. For a five-class dictionary of head tilts, random forest again had the best performance with a test accuracy of ~83% for the short sensor with low placement and ~91% for the long sensor.

Movements farther from the neck were also successfully detected and classified. Sensor data captured from the neck was able to differentiate speaking from static breathing, with ~83% accuracy. The presence and the number of mouth movements was classified with ~68% accuracy. Speech classification was more challenging, achieving up to 62.5% accuracy in differentiating spoken sentences from a four-class dictionary.

7. Conclusions

In this work, we show that subtle neck tilts, mouth movements, and speech can be detected and classified using an inexpensive flex sensor placed at the neck, and thus can prove to be enabling technology for use in software interfaces.

A flex sensor incorporated into a shirt collar or as part of a necklace opens new possibilities for software interaction. The accuracy of the classification of head tilts and their socially undistruptive nature makes head tilting a good option for signally software micro-interactions. For example, a tilt of the head can dismiss a smartwatch notification.

As head gestures can be made during the course of natural speech, the detection of speech and mouth movements allows for the interface to be tailored to times when a person is not speaking and thus improve the interface with greater context awareness.

Author Contributions: Conceptualization, A.N.; methodology, J.L., P.I., K.M. and A.N.; software, J.L. and P.I.; validation, J.L. and P.I.; investigation, J.L., P.I., K.M. and A.N.; writing—original draft preparation, J.L., P.I. and A.N.; writing—review and editing, J.L., A.N. and K.M.; visualization, J.L., P.I. and A.N.; supervision, A.N. and K.M.; project administration, A.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by CSUN Research, Scholarship, and Creative Activity (RSCA) 2021–2022, PI: Ani Nahapetian.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Orphanides, A.K.; Nam, C.S. Touchscreen interfaces in context: A systematic review of research into touchscreens across settings, populations, and implementations. *Appl. Ergon.* **2017**, *61*, 116–143. [[CrossRef](#)] [[PubMed](#)]
2. Jonsson, P.; Johnson, P.W.; Hagberg, M.; Forsman, M. Thumb joint movement and muscular activity during mobile phone texting—A methodological study. *J. Electromyogr. Kinesiol.* **2011**, *21*, 363–370. [[CrossRef](#)]
3. Lee, M.; Hong, Y.; Lee, S.; Won, J.; Yang, J.; Park, S.; Chang, K.-T.; Hong, Y. The effects of smartphone use on upper extremity muscle activity and pain threshold. *J. Phys. Ther. Sci.* **2015**, *27*, 1743–1745. [[CrossRef](#)] [[PubMed](#)]
4. Dannenberg, R.B.; Amon, D. A gesture based user interface prototyping system. In Proceedings of the 2nd Annual ACM SIGGRAPH Symposium on User Interface Software and Technology (UIST '89), Williamsburg, VA, USA, 13–15 November 1989; Association for Computing Machinery: New York, NY, USA, 1989; pp. 127–132. [[CrossRef](#)]
5. McGuckin, S.; Chowdhury, S.; Mackenzie, L. Tap 'n' shake: Gesture-based smartwatch-smartphone communications system. In Proceedings of the 28th Australian Conference on Computer-Human Interaction (OzCHI '16), Launceston, TAS, Australia, 29 November–2 December 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 442–446. [[CrossRef](#)]
6. Deponti, D.; Maggiorini, D.; Palazzi, C.E. Smartphone's psychiatric serious game. In Proceedings of the 2011 IEEE 1st International Conference on Serious Games and Applications for Health (SEGAH '11), Braga, Portugal, 9–11 November 2011; pp. 1–8.
7. Deponti, D.; Maggiorini, D.; Palazzi, C.E. DroidGlove: An android-based application for wrist rehabilitation. In Proceedings of the 2009 IEEE International Conference on Ultra Modern Telecommunications (ICUMT 2009), St. Petersburg, Russia, 12–14 October 2009; pp. 1–7. [[CrossRef](#)]
8. Moazen, D.; Sajjadi, S.A.; Nahapetian, A. AirDraw: Leveraging smart watch motion sensors for mobile human computer interactions. In Proceedings of the 2016 13th IEEE Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 9–12 January 2016; pp. 442–446. [[CrossRef](#)]
9. Vance, E.; Nahapetian, A. Bluetooth-based context modeling. In Proceedings of the 4th ACM MobiHoc Workshop on Experiences with the Design and Implementation of Smart Objects (SMARTOBJECTS '18), Los Angeles, CA, USA, 25 June 2018; Association for Computing Machinery: New York, NY, USA, 2018; pp. 1–6. [[CrossRef](#)]
10. Holmes, A.; Desai, S.; Nahapetian, A. LuxLeak: Capturing computing activity using smart device ambient light sensors. In Proceedings of the 2nd Workshop on Experiences in the Design and Implementation of Smart Objects (SmartObjects '16), New York City, NY, USA, 3–7 October 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 47–52. [[CrossRef](#)]
11. Papisyan, A.; Nahapetian, A. LightVest: A wearable body position monitor using ambient and infrared light. In Proceedings of the 9th International Conference on Body Area Networks (BodyNets '14), London, UK, 29 September–1 October 2014; ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering): Brussels, Belgium, 2014; pp. 186–192. [[CrossRef](#)]
12. Velloso, E.; Schmidt, D.; Alexander, J.; Gellersen, H.; Bulling, A. The Feet in Human–Computer Interaction. *ACM Comput. Surv.* **2015**, *48*, 1–35. [[CrossRef](#)]
13. Scott, J.; Dearman, D.; Yatani, K.; Truong, K.N. Sensing foot gestures from the pocket. In Proceedings of the 23rd Annual ACM Symposium on User Interface Software and Technology (UIST '10), New York City, NY, USA, 3–6 October 2010; Association for Computing Machinery: New York, NY, USA, 2010; pp. 199–208. [[CrossRef](#)]
14. Ohn-Bar, E.; Tran, C.; Trivedi, M. Hand gesture-based visual user interface for infotainment. In Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '12), Portsmouth, NH, USA, 17–19 October 2012; Association for Computing Machinery: New York, NY, USA, 2012; pp. 111–115. [[CrossRef](#)]
15. Davis, J.W.; Vaks, S. A perceptual user interface for recognizing head gesture acknowledgements. In Proceedings of the 2001 Workshop on Perceptive user Interfaces (PUI '01), Orlando, FL, USA, 15–16 November 2001; Association for Computing Machinery: New York, NY, USA, 2001; pp. 1–7. [[CrossRef](#)]
16. Li, H.; Trutoiu, L.C.; Olszewski, K.; Wei, L.; Trutna, T.; Hsieh, P.-L.; Nicholls, A.; Ma, C. Facial performance sensing head-mounted display. *ACM Trans. Graph.* **2015**, *34*, 1–9. [[CrossRef](#)]

17. McNamara, A.; Kabeerdoss, C.; Egan, C. Mobile User Interfaces based on User Attention. In Proceedings of the 2015 Workshop on Future Mobile User Interfaces (FutureMobileUI '15), Florence, Italy, 19–22 May 2015; Association for Computing Machinery: New York, NY, USA, 2015; pp. 1–3. [[CrossRef](#)]
18. Yatani, K.; Truong, K.N. BodyScope: A wearable acoustic sensor for activity recognition. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12), Pittsburgh, Pennsylvania, 5–8 September 2012; Association for Computing Machinery: New York, NY, USA, 2012; pp. 341–350. [[CrossRef](#)]
19. Bi, Y.; Xu, W.; Guan, N.; Wei, Y.; Yi, W. Pervasive eating habits monitoring and recognition through a wearable acoustic sensor. In Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth '14), Oldenburg, Germany, 20–23 May 2014; ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering): Brussels, Belgium, 2014; pp. 174–177. [[CrossRef](#)]
20. Erzin, E. Improving Throat Microphone Speech Recognition by Joint Analysis of Throat and Acoustic Microphone Recordings. *IEEE Trans. Audio Speech, Lang. Process.* **2009**, *17*, 1316–1324. [[CrossRef](#)]
21. Turan, M.T.; Erzin, E. Empirical Mode Decomposition of Throat Microphone Recordings for Intake Classification. In Proceedings of the 2nd International Workshop on Multimedia for Personal Health and Health Care (MMHealth '17), Mountain View, CA, USA, 23 October 2017; Association for Computing Machinery: New York, NY, USA, 2017; pp. 45–52. [[CrossRef](#)]
22. Cohen, E.; Stogin, W.; Kalantarian, H.; Pfammatter, A.F.; Spring, B.; Alshurafa, N. SmartNecklace: Designing a wearable multi-sensor system for smart eating detection. In Proceedings of the 11th EAI International Conference on Body Area Networks (BodyNets '16), Turin, Italy, 15–16 December 2016; ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering): Brussels, Belgium, 2016; pp. 33–37.
23. Hirsch, M.; Cheng, J.; Reiss, A.; Sundholm, M.; Lukowicz, P.; Amft, O. Hands-free gesture control with a capacitive textile neckband. In Proceedings of the 2014 ACM International Symposium on Wearable Computers (ISWC '14), Seattle, WA, USA, 13–17 September 2014; Association for Computing Machinery: New York, NY, USA, 2014; pp. 55–58. [[CrossRef](#)]
24. Cheng, J.; Zhou, B.; Kunze, K.; Rheinländer, C.C.; Wille, S.; Wehn, N.; Weppner, J.; Lukowicz, P. Activity recognition and nutrition monitoring in every day situations with a textile capacitive neckband. In Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication (UbiComp '13 Adjunct), Zurich, Switzerland, 8–12 September 2013; Association for Computing Machinery: New York, NY, USA, 2013; pp. 155–158. [[CrossRef](#)]
25. Kalantarian, H.; Alshurafa, N.; Sarrafzadeh, M. A Wearable Nutrition Monitoring System. In Proceedings of the 2014 11th International Conference on Wearable and Implantable Body Sensor Networks (BSN '14), Zurich, Switzerland, 16–19 June 2014; IEEE Computer Society: Washington, DC, USA, 2014; pp. 75–80. [[CrossRef](#)]
26. Kalantarian, H.; Motamed, B.; Alshurafa, N.; Sarrafzadeh, M. A wearable sensor system for medication adherence prediction. *Artif. Intell. Med.* **2016**, *69*, 43–52. [[CrossRef](#)] [[PubMed](#)]
27. Morrow, K.; Wilbern, D.; Taghavi, R.; Ziat, M. The effects of duration and frequency on the perception of vibrotactile stimulation on the neck. In Proceedings of the 2016 IEEE Haptics Symposium (HAPTICS '16), Philadelphia, PA, USA, 8–11 April 2016; pp. 41–46.
28. Yamazaki, Y.; Hasegawa, S.; Mitake, H.; Shirai, A. Neck strap haptics: An algorithm for non-visible VR information using haptic perception on the neck. In Proceedings of the ACM SIGGRAPH 2019 Posters (SIGGRAPH '19), Los Angeles, CA, USA, 28 July 2019; Association for Computing Machinery: New York, NY, USA, 2019. Article 60, pp. 1–2. [[CrossRef](#)]
29. Yamazaki, Y.; Mitake, H.; Hasegawa, S. Tension-based wearable vibroacoustic device for music appreciation. In Proceedings of the International Conference on Human Haptic Sensing and Touch Enabled Computer Applications (EuroHaptics '16), London, UK, 4–7 July 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 273–283.
30. Ephrat, A.; Peleg, S. Vid2speech: Speech reconstruction from silent video. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '17), New Orleans, LA, USA, 5–9 March 2017; pp. 5095–5099.
31. Ephrat, A.; Halperin, T.; Peleg, S. Improved speech reconstruction from silent video. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 455–462.
32. Kumar, Y.; Jain, R.; Salik, M.; Shah, R.; Zimmermann, R.; Yin, Y. MyLipper: A Personalized System for Speech Reconstruction using Multi-view Visual Feeds. In Proceedings of the 2018 IEEE International Symposium on Multimedia (ISM), Taichung, Taiwan, 10–12 December 2018; pp. 159–166. [[CrossRef](#)]
33. Kimura, N.; Hayashi, K.; Rekimoto, J. TieLent: A Casual Neck-Mounted Mouth Capturing Device for Silent Speech Interaction. In Proceedings of the International Conference on Advanced Visual Interfaces, Salerno, Italy, 28 September–2 October 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 1–8. [[CrossRef](#)]
34. Sun, K.; Yu, C.; Shi, W.; Liu, L.; Shi, Y. Lip-Interact: Improving Mobile Device Interaction with Silent Speech Commands. In Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18), Berlin, Germany, 14–17 October 2018; Association for Computing Machinery: New York, NY, USA, 2018; pp. 581–593. [[CrossRef](#)]
35. Manabe, H.; Hiraiwa, A.; Sugimura, T. Unvoiced speech recognition using EMG—Mime speech recognition. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems (CHI EA '03)*; Association for Computing Machinery: New York, NY, USA, 2003; pp. 794–795. [[CrossRef](#)]
36. Maier-Hein, L.; Metze, F.; Schultz, T.; Waibel, A. Session independent non-audible speech recognition using surface electromyography. In Proceedings of the IEEE Workshop on Automatic Speech Recognition and Understanding, Cancun, Mexico, 27 November–1 December 2005; pp. 331–336. [[CrossRef](#)]

37. Sahni, H.; Bedri, A.; Reyes, G.; Thukral, P.; Guo, Z.; Starner, T.; Ghovanloo, M. The tongue and ear interface: A wearable system for silent speech recognition. In Proceedings of the 2014 ACM International Symposium on Wearable Computers (ISWC '14), Seattle, WA, USA, 13–17 September 2014; Association for Computing Machinery: New York, NY, USA, 2014; pp. 47–54. [[CrossRef](#)]
38. Li, R.; Wu, J.; Starner, T. TongueBoard: An Oral Interface for Subtle Input. In Proceedings of the 10th Augmented Human International Conference 2019 (AH2019), Reims, France, 11–12 March 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 1–9. [[CrossRef](#)]
39. Zhang, Q.; Gollakota, S.; Taskar, B.; Rao, R.P. Non-intrusive tongue machine interface. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14), Toronto, ON, Canada, 26 April–1 May 2014; Association for Computing Machinery: New York, NY, USA, 2014; pp. 2555–2558. [[CrossRef](#)]
40. Nguyen, P.; Bui, N.; Nguyen, A.; Truong, H.; Suresh, A.; Whitlock, M.; Pham, D.; Dinh, T.; Vu, T. TYTH-Typing on Your Teeth: Tongue-Teeth Localization for Human-Computer Interface. In Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '18), Munich, Germany, 10–15 June 2018; Association for Computing Machinery: New York, NY, USA, 2018; pp. 269–282. [[CrossRef](#)]
41. Ashbrook, D.; Tejada, C.; Mehta, D.; Jiminez, A.; Muralitharam, G.; Gajendra, S.; Tallents, R. Bitey: An exploration of tooth click gestures for hands-free user interface control. In Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '16), Florence, Italy, 6–9 September 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 158–169. [[CrossRef](#)]
42. Crossan, A.; McGill, M.; Brewster, S.; Murray-Smith, R. Head tilting for interaction in mobile contexts. In Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '09), Bonn, Germany, 15–18 September 2009; Association for Computing Machinery: New York, NY, USA, 2009; pp. 1–10. [[CrossRef](#)]
43. LoPresti, E.; Brienza, D.M.; Angelo, J.; Gilbertson, L.; Sakai, J. Neck range of motion and use of computer head controls. In Proceedings of the Fourth International ACM Conference on Assistive Technologies (Assets '00), Arlington, VA, USA, 13–15 November 2000; Association for Computing Machinery: New York, NY, USA, 2000; pp. 121–128. [[CrossRef](#)]
44. Ando, T.; Kubo, Y.; Shizuki, B.; Takahashi, S. CanalSense: Face-Related Movement Recognition System based on Sensing Air Pressure in Ear Canals. In Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17), Quebec City, QC, Canada, 22–25 October 2017; Association for Computing Machinery: New York, NY, USA, 2017; pp. 679–689. [[CrossRef](#)]
45. Lacanlale, J.; Isayan, P.; Mkrtchyan, K.; Nahapetian, A. Look Ma, No Hands: A Wearable Neck-Mounted Interface. In Proceedings of the Conference on Information Technology for Social Good (GoodIT '21), Rome, Italy, 9–11 September 2021; Association for Computing Machinery: New York, NY, USA, 2021; pp. 13–18. [[CrossRef](#)]