

# In-Mouth Robotic Bite Transfer with Visual and Haptic Sensing

Lorenzo Shaikewitz<sup>1,\*</sup>, Yilin Wu<sup>2,\*</sup>, Suneel Belkhale<sup>2,\*</sup>, Jennifer Grannen<sup>2</sup>, Priya Sundaesan<sup>2</sup>, Dorsa Sadigh<sup>2</sup>

California Institute of Technology<sup>1</sup>, Stanford University<sup>2</sup>. \* denotes equal contribution.

**Abstract**—Assistance during eating is essential for those with severe mobility issues or eating risks. However, dependence on traditional human caregivers is linked to malnutrition, weight loss, and low self-esteem. For those who require eating assistance, a semi-autonomous robotic platform can provide independence and a healthier lifestyle. We demonstrate an essential capability of this platform: safe, comfortable, and effective transfer of a bite-sized food item from a utensil directly to the inside of a person’s mouth. Our system uses a force-reactive controller to safely accommodate the user’s motions throughout the transfer, allowing full reactivity until bite detection then reducing reactivity in the direction of exit. Additionally, we introduce a novel dexterous wrist-like end effector capable of small, unimposing movements to reduce user discomfort. We conduct a user study with 11 participants covering 8 diverse food categories to evaluate our system end-to-end, and we find that users strongly prefer our method to a wide range of baselines. Appendices and videos are available at our website: <https://tinyurl.com/btICRA>.

**Index Terms**—Physical Human Robot Interaction, Physically Assistive Devices, Human Robot Collaboration

## I. INTRODUCTION

Robots have the potential to assist people in many aspects of their daily lives. Eating is an integral part of our day, and yet millions of people, from young children to hospitalized patients and the elderly, require assistance from caregivers to perform this essential function. This includes nearly 70% of hospitalized elderly and at least a quarter of nursing home patients [1]. However, many of those who benefit from assistance suffer from a loss of independence that can reduce their sense of self-worth, and may lead to malnutrition and weight loss [1]. For these individuals, robotic solutions that automate feeding promise a return to independence and a healthier lifestyle.

One critical piece of the feeding process is robot-assisted *bite transfer*: bringing various food items into a user’s mouth using a tool, allowing them to safely take a bite, and then removing the tool from the mouth (see Figure 1). While several works have studied fork-based bite transfer, most works bring the food item to a point outside the user’s mouth and then have the user move forward to bite the food on their own [2]. This is simply not practical for the 300,000 people in the U.S. alone who suffer from spinal cord injuries and *require* caregivers to feed them, nor is it comfortable even for the broader population of people with motor impairment disabilities [3]. The *in-mouth* phase of bite transfer is less studied, and it is particularly challenging since we are bringing a tool (often a fork) inside one of the most fragile and sensitive parts of the human body. It is also difficult to model the human mouth and

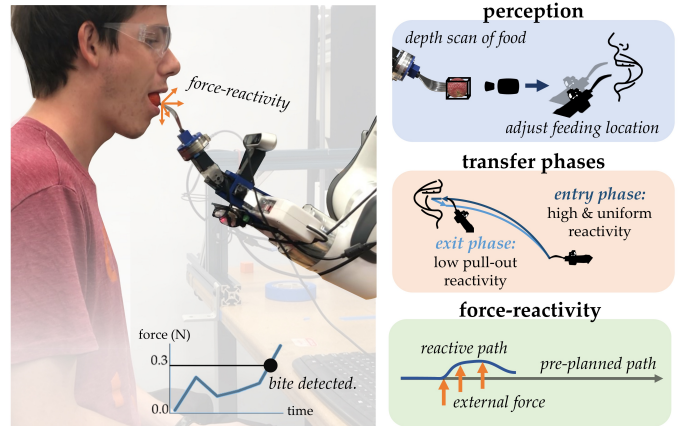


Fig. 1: **Robust Robotic Bite Transfer.** Our system safely and comfortably feeds a user bite-sized food items. For an acquired food item, the system detects the user’s mouth and adjusts its target position based on a depth scan of the food. To feed the user, it follows an arced trajectory and monitors force to detect a bite. Throughout feeding, it employs a force-reactive controller for user comfort and safety.

its actuation, nor do we know when or how the human wants to take a bite. Furthermore, the robot needs to be able to handle a wide diversity of foods and food orientations on the tool. Most recently, [Belkhale et al.](#) developed an in-mouth transfer system which balances comfort and bite efficiency, where comfort was defined as encroaching on a user’s personal space when approaching the mouth. They utilize a single reactive controller for adapting the end effector pose to the forces the user applies on the utensil. While this method improves comfort for transfer when *approaching* the mouth, this work does not address user comfort during the in-mouth transfer process.

There are several key ways in which prior work fails to address comfort during in-mouth transfer. Firstly, prior work uses a single, unchanging reactive controller for the entire bite transfer process. It is important to be reactive to the person being fed; however, the type of reactivity changes based on the *phase* of bite transfer. For example, when feeding a strawberry to a user, we might want to be very compliant during the entry into the mouth, stiff during the bite, and then stiff in the direction of exit in case the user wants to pull the strawberry off the fork. A single reactive controller would fail to take this dimension of comfort into account. Secondly, prior works directly mount tools to the robot end effector for transfer, which can force the robot to make large joint position movements just to adjust the tool orientation. Often, this motion at robot joints early in the kinematic chain leads to large movements during transfer, which is jarring and

uncomfortable for users [4]. This is analogous to a caregiver reorienting their elbow, shoulder, and even torso while they are feeding someone, as opposed to subtle movements at the wrist. Thirdly, the prior in-mouth bite transfer work evaluates their method only on carrots [4]. It is necessary to test on more diverse and complex food items (e.g., deformable foods like cheesecake, irregularly shaped foods like broccoli) to draw meaningful conclusions about comfort for in-mouth transfer.

In this work, we develop a comfortable system for in-mouth bite transfer across a diverse set of bite-sized foods. To enable comfortable in-mouth transfer of diverse food items, this work builds on two key insights. First, we posit that bite transfer can be split into two phases: mouth entry and mouth exit, separated by bite detection. We can then design *phase-specific* reactive controllers enabling comfortable physical interactions with the robot during the in-mouth transfer. During mouth entry, our system obtains precise measurements of the food item and the user’s facial keypoints using vision, and then utilizes this information to bring the food item into the user’s mouth. During entry and exit, the system uses a reactive controller to avoid sensitive collisions with the user’s mouth, where the specific reactivity coefficients can be different for each phase. Second, inspired by the mobility of the human wrist, we aim to enable more degrees of freedom at the end effector to reduce the jarring motion caused by robot joints earlier in the kinematic chain. To accomplish this, we develop a custom “wrist” joint to enable fine-grained reactivity while in the mouth. Using this full system, we conduct a user study with 11 participants and 8 different bite-sized food items to evaluate our system. Our method is substantially preferred over the baselines, and we perform an in-depth analysis of qualitative and quantitative user feedback to guide future work.

## II. RELATED WORK

### A. Robot-Assisted Feeding

Robot-assisted feeding has become an active area of research and development in recent years. There are several robotic feeding systems on the market currently; however, they lack widespread adoption due to minimal autonomy or dependence on food preparation beforehand [5, 6, 7, 8, 9]. In the literature, the autonomous feeding problem is generally split into two phases: bite acquisition and bite transfer.

**Bite Acquisition:** Within bite acquisition, Feng et al. developed the SPANet network, which uses supervised learning to map segmented foods of known classes to a corresponding set of discrete actions [10]. Gordon et al. added an online contextual bandits framework for updating the SPANet action prediction using haptic information [11]. Sundaresan et al. introduce a probing action during acquisition to zero-shot predict the optimal acquisition action given both visual and haptic information, without the need for multiple trials, across a wide range of foods [12]. Beyond fork-based acquisition, Grannen et al. develop a spoon-based acquisition framework that leverages two robot arms to provide stability during scooping [13]. However, none of these approaches consider the problem of bite transfer once the food is acquired. Recently, Gallenberger et al. demonstrate that bite transfer indeed de-

pends on bite acquisition, but they do not provide an approach for feeding bites of food in a general fashion [14].

**Bite Transfer:** Similar to acquisition, autonomous systems for bite transfer usually rely on discrete food classes with predetermined feeding trajectories [10, 14, 15, 16]. These works also generally bypass in-mouth transfer by navigating the utensil to a predetermined pose outside of the user’s mouth, requiring the user to reach the utensil on their own to complete the feeding process. One recent work demonstrated an in-mouth transfer system that optimized for comfort and efficiency during planning; however, this work primarily studied comfort during the *approach* trajectory and did not consider the in-mouth comfort [4]. These works also utilize rigidly mounted utensils for feeding, leading to jarring motion of the robot during transfer as described in Sec. I. Each of these prior works utilizes a limited set of foods for bite transfer evaluation. Our work aims to specifically improve user comfort during in-mouth transfer.

### B. Physical Human-Robot Interaction

Physical human-robot interaction encompasses scenarios where the human and robot must physically interact in the real world in a safe manner. Prior work studies how robots can safely respond to such interaction [17, 18, 19], including the use of reactive controllers [20, 21]. Several works look at how robots can learn from physical interaction, for example through physical corrections [22, 23]. Bite transfer is a specific type of physical human robot interaction. However, bite transfer involves safety and comfort in a constrained and sensitive workspace, a human’s mouth, where it is critical for the robot to be precise and reactive to the user. Most closely related to the bite transfer problem is robot-human handovers, which usually involve the passing of some item from a robot arm to a human hand. These works do not usually consider using an intermediate tool for handover [24, 25, 26]. Unlike in the feeding domain, handovers usually involve a person’s hand as the recipient. Cakmak et al. find clear human preferences for object orientations and grasp placement in the application-agnostic handover problem. However, passing an item to a person’s mouth presents a host of new challenges involving safety and comfort. Canal et al. show the relevance of individual preferences to the bite transfer process. In our work, we consider the in-mouth handover of bite-sized food items, and we study how to make such transfer both safe and comfortable.

## III. IN-MOUTH BITE TRANSFER SYSTEM

This section details the operation and design of our robotic bite transfer system. We begin with a description of the hardware platform in Section III-A, describing the general purpose robotic arm and our custom-built actuated wrist-like end effector. Then, in Section III-B, we present our phased force-reactive controller, which provides a basic measure of safety and enhances comfort by adjusting during the feeding process. Next, Sections III-C and III-D detail the mechanisms of the bite transfer system. In Section III-A, we explain how the system chooses a target location using a depth camera image and face detection. The arced trajectory to reach this target and

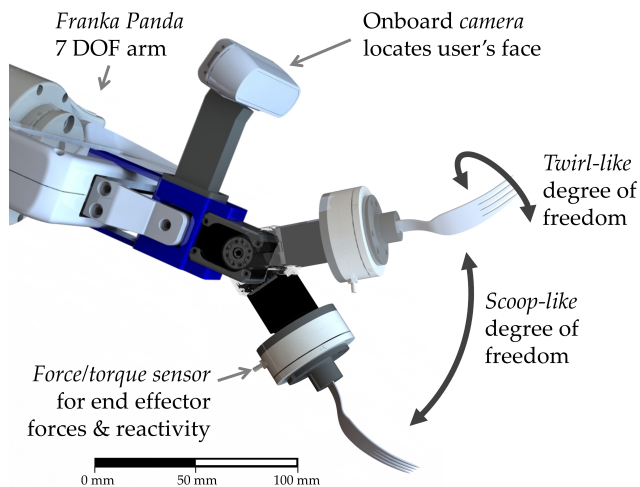


Fig. 2: **A Wrist-Like End Effector.** We augment the Franka Panda’s gripper with a 2 degree-of-freedom (DOF) end effector designed for feeding motions. This allows simple scooping and twirling motions without moving the entire arm. The wrist also contains perception hardware (mounted RGB-D camera and force/torque sensor) for adjusting to the user’s movements.

the bite detection system are described in Section III-D. By combining a dexterous end effector and force-reactive control with a robust targeting and transfer procedure, our system safely navigates into the user’s mouth and withdraws upon a bite. An overview of this system is shown in Figure 1.

#### A. Actuated Wrist-Like End Effector

Our system relies on a 7 degree-of-freedom general purpose manipulator (Franka Emika Panda) gripping a custom-built actuated fork assembly. The fork assembly, shown in detail in Figure 2, adds two degrees of freedom designed to mimic the primary movements of a wrist: twirl-like rotation about the fork’s long axis, and rotation about a single perpendicular axis to allow a scooping motion. This extends the range of motion of the robot and increases its null space. For feeding, these additions allow rapid adjustment of the fork’s position without significant movement of the Panda arm. Fewer movements may also increase feelings of comfort around the arm.

The fork assembly (Figure 2) is composed of two servo motors (Dynamixel XC330-M288-T) connected in series and plastic mounting hardware. It is connected to the Panda’s gripper using a 3D printed mount, which also holds a camera (Intel RealSense D435). The twirling degree of freedom is located after the scoop-like degree of freedom at the end of the kinematic chain. A force/torque sensor (ATI Mini45 6-axis F/T sensor) is located between the fork and the second servo motor, forming the tip of the end effector.

#### B. Force-Reactive Controller

Our wrist end effector provides a hardware platform for feeding, but bringing a fork into a user’s mouth requires careful consideration of safety. In addition to typical safety features like an e-stop and slow movement, our system is responsive to forces at the tool using a force-reactive controller.

In traditional impedance control, the desired position and velocity is used to make a desired end effector wrench,  $f$ , which coupled with the Jacobian  $J$  gives necessary joint

torques  $\tau$ . Our controller adds reactivity to the forces on the tool, which is equivalent to adjusting the desired wrench before it is passed through the impedance control relation. Specifically, we add an external force to impedance control:  $\tau = J^T (f - \bar{f}_m)$ , where  $\bar{f}_m$  comes from a proportional-integral (PI) term on the measured tool forces  $f_m$ . The PI term is given by  $\bar{f}_m = k_P f_m + k_I \int f_m dt$ .

This added reactivity operates at a 1kHz control rate and enables compliance to fine-grained user movements. We tune the sensitivity of this system by adjusting the vector proportionality constants  $k_P$  and  $k_I$ . To prevent unexpected movements from the robot, we set the torque components of these constants to zero. Note that the force reactivity does not extend to the wrist end effector’s degrees of freedom, so using torque reactivity requires larger movements.

**Phased Reactivity.** The sensitivity of the system is key to the safety and comfort of the bite transfer; however, it is not necessary for the reactivity to remain the same in all axes throughout the transfer. Our method incorporates a phased force-reactive controller, using one set of constants as the fork enters the mouth (the *entrance phase*) and a different set of constants as the fork is leaving the mouth (the *exit phase*), as shown in Figure 1. In the former phase, users benefit from high reactivity because they may safely guide the fork to a comfortable transfer location. Thus, we keep  $k_P$  and  $k_I$  relatively high ( $k_P = 7$  and  $k_I = 20$  Hz for all forces, and  $k_P = k_I = 0$  for all torques). During the exit phase, high reactivity can inhibit transfer as the force of a bite prevents successful withdrawal. Therefore, we reduce the reactive constants only in the exit direction ( $k_P = 2$  and  $k_I = 1$  Hz for forces into or out of the mouth). This maintains compliance in the directions orthogonal to the exit direction while combating the resistance of the bite for a successful withdrawal. In addition, the reactive constants may be tuned to each individual’s preference, which we leave to future work.

#### C. Defining a Target Location

As the system is designed to operate with any bite sized food item, we adjust the location where the fork enters the mouth only according to the food’s placement on the fork. This is a two phase process: first, an external camera uses a depth scan of the food item to determine its bounding box relative to the fork. Then, a camera mounted on the end effector identifies the location of the user’s mouth and applies the offsets computed from the depth scan’s bounding box (Figure 3).

**Offset Generation.** In the depth scan phase, the tines of the fork are facing an external fixed short-range depth camera (Intel RealSense D405). This perspective is ideal for getting offsets in the plane parallel to the user’s face; the camera views the end of the fork as if the fork was moving towards it. The resulting point cloud, obtained at 0.1 mm resolution, is transformed into a “mouth frame” (Figure 3) located at the tip of the fork. This corresponds to a point just below the teeth in the open mouth: a safe feeding location if there were no food on the fork, and must be adjusted based on food orientation. Specifically, this mouth frame has its  $z$ -axis pointing out of the

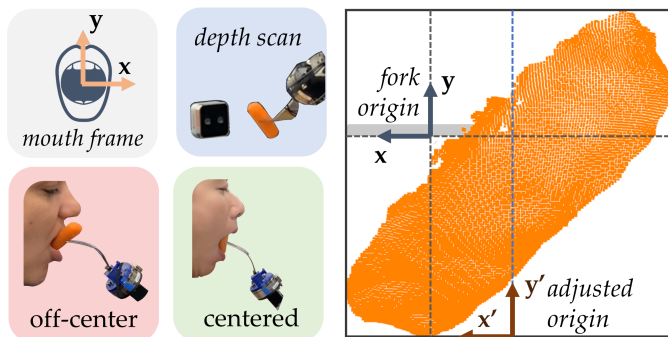


Fig. 3: **Adjusting the Target Position.** For any placement of a bite-sized food item on the end of the fork, we compute a safe target position given a depth scan taken from the mouth perspective and centered at optimal entry coordinates for the fork alone. Our system uses the scanned point cloud to recenter and lower the food by averaging the bounding box  $x$  coordinates and adjusting by the minimum  $y$  coordinate, respectively.

mouth and  $y$ -axis pointing upwards towards the user's eyes. The  $x$ -axis is along the lip, pointing to the user's left.

From this coordinate system, the axis-aligned bounding box of the point cloud gives the necessary offsets in the face ( $x$ - $y$ ) plane. The  $x$ -axis offset (along the lips) aligns the center of the food, rather than the fork, with the mouth center. This is especially helpful for large, off-centered acquisitions, like the one shown in Figure 3. Assuming the lower jaw does most of the movement to open the mouth, the  $y$  direction offset moves the top of the food down to a position where it will not contact the upper teeth. We focus on bite-sized foods in this work so we do not perform a depth measurement to adjust the distance into the mouth that the fork enters. Instead, the distance is fixed to 18 mm, which we empirically found to be comfortable and safe for most users.

**Face Detection.** These offsets are added to location of the mouth, which is obtained using a pretrained Histogram of Oriented Gradients (HOG) model that is publicly available through DLib [29]. After the scanning phase is complete, users are instructed to remain still with their mouth closed while the network obtains a mouth center point. This center point is adjusted according to the previously determined offsets to compute the target feeding position. Since the network cannot recognize faces when it gets too close to the user, we rely solely on the initial scan to determine the mouth position; users are expected to remain still throughout the 10 second trajectory. This assumption, while slightly prohibitive for able-bodied users, is reasonable for many individuals who require feeding assistance but have difficulty moving their heads.

#### D. Trajectory and Transfer

Throughout the bite transfer process, the trajectory of the robot is designed to maximize comfort by emulating the ways humans feed themselves. After offset and face detection, the end effector moves to a low position and the fork is rotated to an upside down position for easier transfer from below. Then, it follows an arced trajectory (radius 0.45 m, similar to the length of a lower arm, and centered directly below the target position near the elbow) to a point just in front of the mouth with the fork angled upward. This emulates the motion of an elbow and brings the food in from below, keeping the bulk

of the robot arm away from the user. Then, the robot linearly enters the mouth at its target position and slightly lowers the fork to keep it from hitting the top of the user's mouth.

**Transfer Phases.** Based on the different levels of sensitivity, the bite transfer is divided into two phases, i.e., the entry phase and the exit phase. The system can automatically divide the two phases with a bite detection function. The bite is detected when the magnitude of the force from  $y$  axis is greater than the threshold of 0.3 N (obtained in early bite testing with our system). During the entry phase, the end effector remains motionless and reactive as the system waits for a bite. If no bite is detected within a wait period of 1.5 seconds, the bite detection times out and begins the exit phase. This relatively short timeout allows the user to reject the food or initiate transfer without a forceful bite. In the exit phase, the fork simply leaves the mouth in a linear trajectory and then follows the same arc path down to its starting position.

The result is a system capable of robustly feeding a variety of food items directly into a user's mouth. Sampled real-world trajectories are shown in Appendix A<sup>1</sup>.

## IV. EXPERIMENTAL EVALUATION

To evaluate the safety, comfort, and overall experience of using our system, we conducted a user study feeding 11 participants food items with our bite transfer system and 4 baselines. We additionally show that our custom wrist end-effector can improve upon several dimensions of user comfort.

### A. Experimental Setup

**Participants and Procedure.** Our study included 11 able-bodied<sup>2</sup> participants and 8 varied food items. Participants (6 female, 5 male, aged between 22 and 57) were somewhat experienced with robots, with 6/11 indicating they had interacted with one before. 7 participants reported they had never been fed as an adult, and 3 reported never having fed someone else. We used a within-subjects design, feeding each subject randomly chosen three food items using each of five methods, twice each. Additionally, we gave each subject one additional feeding trial per method. This study was approved by the Institutional Review Board of Stanford University.

**Safety Considerations.** Safety of participants was a primary objective. At each major planning phase (scanning and face detection), the system paused for proctor confirmation before moving towards the user. On top of the compliant reactive controller, we included a software stop: motion was automatically aborted if the end effector force exceeded 3N, a mild amount of contact. Additionally, an emergency stop button was in place and available to both the proctor and participant.

### B. Study Design

**Independent Variables: Food Settings.** We evaluate the bite transfer system with 8 food item classes: carrot, strawberry, blueberry, pineapple, cherry tomato, broccoli, cheesecake, and tofu. These food classes cover a wide spectrum of sizes,

<sup>1</sup>The appendix may be found at our website: <https://tinyurl.com/bticra>

<sup>2</sup>Able-bodied participants were chosen to protect vulnerable individuals amid the ongoing COVID-19 pandemic.

shapes, and deformabilities (see Table X in Appendix B). All food items were cut to bite-size, and we gave participants the opportunity to opt out of eating specific foods.

**Independent Variables: Feeding Method.** We compare our method against three robotic feeding baselines and one human oracle baseline. *Fixed Pose* is from prior work and is not an in-mouth bite-transfer method [14]. Instead, the robot moves the fork to a position outside the user’s mouth and allows the user to take a bite on their own. *Less Reactive* and *More Reactive* are two ablations of our system with less and more force-reactivity respectively. Specifically, the less reactive method uses  $k_P = 2$  and  $k_I = 2$  Hz for all forces throughout feeding, while the more reactive method uses  $k_P = 10$  and  $k_I = 30$  Hz for all forces. We note that *More Reactive* is representative of the in-mouth bite transfer method in Belkhale et al., which uses a single, highly reactive controller. *Human Oracle* involves the user being fed by another human, which serves as a reasonable upper bound for our robotic feeding system.

**Dependent Measures.** We consider both subjective and objective measures. After each food item was fed with all methods, participants were asked to rank the methods considering safety, comfort, and overall preference (*Safety*, *Comfort*, and *Overall Ranking*). After the study, participants provided a *Final Ranking* of the methods based only on their overall preference. Per feeding method, we measured the fraction of trials in which the bite is successfully transferred (*Success Rate*).

**Hypothesis.** We hypothesized that:

**H1.** By using a multi-phase reactive controller and a more dexterous wrist end-effector, our method will outperform the three robotic baselines in terms of comfort, safety, and overall rankings for all foods.

**H2.** Our method will have a higher transfer success rate compared to the three robotic baselines for all foods.

### C. Results

We present both subjective and objective metrics from our user study to evaluate our system.

**Subjective Metrics.** We show the overall user rankings of each of the five feeding methods (H1) in Figure 4. Additionally, we present user ranking results on the safety, comfort, and overall quality of each feeding method (H1), grouped by food variability axis (shape, size, and deformability—defined in Appendix B), shown in Figure 5.

Both overall and across food item variabilities, we find that users prefer our in-mouth bite transfer method over all other in-mouth feeding methods (Less Reactive and More Reactive), and that our method performs comparably to a fixed pose transfer method (H1). This trend holds across safety and comfort ratings as well, as users noted the method “had a gentle approach and placement of food” and it was “easy to bite, the robot reacted and the motion away from the mouth was good”. In some instances, our method even approaches a human feeding oracle performance, suggesting that our method consistently performs high-quality bite transfer.

We find that users rank the fixed pose baseline relatively highly as they are able to “be active in the feeding process” and

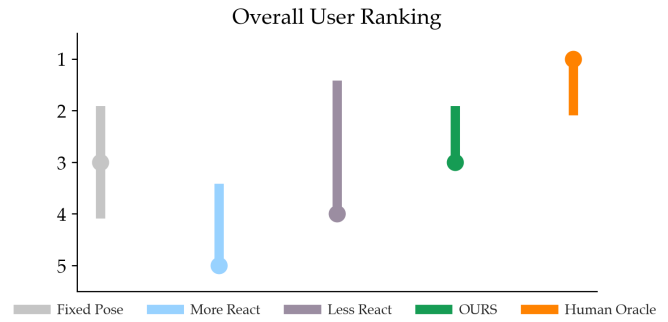


Fig. 4: **Overall User Rankings.** We report the median and interquartile range of user rankings across four robotic feeding methods and one human feeding oracle. We compare our method against two similar baselines with varied reactivity coefficients, and one fixed pose baseline that holds the food outside the mouth and places the burden on the user to move to take a bite. Our in-mouth bite transfer has the highest overall user ranking of the robotic feeding methods, suggesting the effectiveness of our phased reactive method.

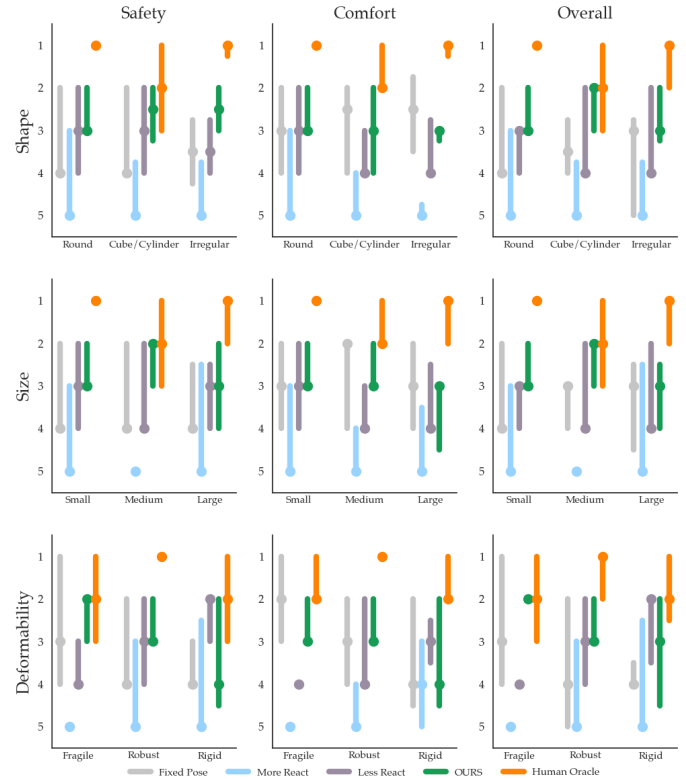


Fig. 5: **User Rankings by Food Property.** We report the median and interquartile range of user rankings of the safety, comfort, and overall quality of the feeding method by food property. As shown in Table X in Appendix B, we group the 8 foods considered (broccoli, blueberry, cherry tomato, strawberry, carrot, pineapple, tofu, and cheesecake) into 3 categories along each of 3 axes of variability: shape, size, and deformability. Each food item appears exactly once in each variability category. One-way ANOVA testing confirms the statistical significance of transfer style (in-mouth vs. out-of-mouth) and reactivity (less, more, ours) on the reported user ratings for comfort alone, and safety, comfort, and overall, respectively. We report additional per food user rankings and exact ranking values in Appendix B.

retain more control over their bite, leading to higher ratings for comfort in particular. This is by nature of the method placing the burden on the user to move their head to take a bite off the fork. While this method may be more comfortable for able-bodied users, moving the mouth to the fork may be impractical for those with mobility issues who stand to benefit the most from an assistive feeding robot system. We report additional

user rankings on head movement in Appendix B.

We note that users prefer our in-mouth bite transfer and consider it to be more safe and comfortable than the other robotic feeding baselines across most food groups; in other words, our method of moderate, phased reactivity is generally preferred. One exception is the rigid food class under food deformability, where the Less Reactive baseline is preferred for safety, comfort, and overall. This is due to the high friction of the rigid food against the fork, making it difficult for the user to pull the food off the fork with a more reactive method such as our system or the More Reactive baseline.

We perform additional one-way ANOVA testing to assess the effects of (1) phased reactivity (*Less Reactive*, *Ours*, or *More Reactive*) and (2) transfer style (in-mouth — *Less Reactive*, *Ours*, and *More Reactive* — or out-of-mouth — *Fixed Pose*) on user ratings. We find that phased reactivity has a statistically significant effect on safety, comfort, and overall ratings, while transfer style has a statistically significant affect on comfort ratings alone (denoted by  $p \leq 0.05$ ), even for a small sample size of 11 users. The results support the hypothesis (H1) that both phased reactivity and transfer style impact a user’s experience. We provide additional details of the ANOVA testing procedure in Appendix B.

Method	Success Rate	Failure Mode		
		Bite Failure	Drop	Imprecise
<i>Fixed Pose</i>	65/66	0	1	0
<i>More Reactive</i>	59/66	6	0	1
<i>Less Reactive</i>	63/66	2	0	1
<i>Ours</i>	62/66	3	1	0
<i>Human Oracle</i>	<b>66/66</b>	0	0	0

TABLE I: **Bite Transfer Success Rate.** We report the success rate and failure modes of all the methods across 66 trials with 8 different foods and 11 users in the experiments. Three types of failure modes happen in the user study. In *Bite Failure*, users fail to take a bite of the food. This can occur when either the user retracts early due to mistrust or fails to take a bite before the fork exits. In *Drop*, the food falls off the fork during the transfer process. *Imprecise* denotes the face detector localizing the mouth inaccurately, causing a near-miss or off-center transfer. The highest success rate is human oracle and the most common failure mode is Bite Failure.

**Objective Metrics.** Our quantitative results measuring bite transfer success (H2) are shown in Table I, grouped by failure mode. Among all the methods, *Human Oracle* represents an upper bound on performance. *Fixed Pose* demonstrates the second highest success as it bypasses in-mouth transfer entirely, and consequently any associated challenges. However, this method is a highly impractical solution for people with mobility impairments because it requires the user to significantly move their head to take a bite. Our proposed approach critically alleviates this need with in-mouth transfer, which naturally introduces more potential for imprecision or bite failure. We note that the *More Reactive* method suffers from the most Bite Failure errors, likely caused by jerkiness in the compliant controller which can be jarring to a user. *Ours* and *Less Reactive* perform comparably and outperform the *More Reactive* method, suggesting the effectiveness of phased transfer and less compliance upon fork exit. Finally, we report the occurrence of food falling off the fork during the transfer process and mouth detection inaccuracies resulting in imprecise transfer. We note that these minor failures are system-level errors, irrespective of the transfer method employed.

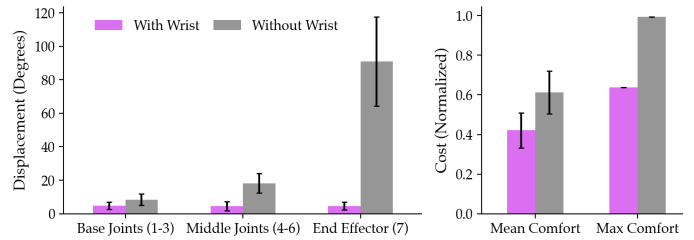


Fig. 6: We generate IK solutions for 10,000 random fork poses using the Wrist joint (With Wrist) as compared to directly mounting the tool (Without Wrist). **Left:** Average Joint Displacement, for the 7 Franka Panda joints (grouped for simplicity). **Right:** Average Comfort Cost, defined as encroaching on the user’s personal space, similarly to prior work [4]. Our wrist mount achieves substantially lower joint displacement (less jarring motion) and comfort cost (less personal space occlusion), at no additional computational cost.

**Wrist Joint Analysis.** In addition to our user study, we analyze and demonstrate the ability of our custom wrist end-effector to improve perceived comfort through simulation experiments in Figure 6. We generate 10,000 sampled tool poses that are likely to occur during feeding, and then compute the inverse kinematics (IK) solution for the robot joint configuration using the damped least squares method. We compute IK across two robot setups, one with the tool mounted on our custom wrist end-effector (9 DOF), and one with the tool directly mounted to the robot’s default end effector (7 DOF). For both setups, we evaluate the average joint displacement across the 7 robot joints for the sampled tool poses. High joint displacement indicates that the joints earlier in the kinematic chain must move more to reach a given tool pose, thus leading to jarring and uncomfortable motion for the user. By adding more degrees of freedom at the end of the kinematic chain, our wrist-like end effector substantially lowers the average joint displacement (see left of Figure 6). Furthermore, we compute a “comfort” cost similar to the one used in Belkhal et al. [4], representing the degree to which the robot encroaches on the user’s personal space. By choosing configurations that stay out of a user’s field of view, the wrist also lowers the average and maximum personal space comfort cost (see right of Figure 6).

## V. CONCLUSION

We present a phased bite transfer system that combines visual and haptic information for comfortable and safe in-mouth bite transfer. Our method uses a custom wrist-like joint for additional degrees of freedom and force-reactivity for increased user comfort during feeding. We present results from a user study with 11 participants comparing our in-mouth bite transfer system to baselines of varied reactivity from prior work, fixed pose feeding methods, and a human feeding oracle. We find that our system is rated the highest of the in-mouth robotic feeding methods both in terms of comfort and safety, and is able to significantly reduce robot joint motions that are often jarring to a user using our novel wrist joint.

In future work, we will explore how to incorporate user feedback and preferences in bite transfer to customize a robotic feeding system to each user. Beyond the transfer component, we will consider different trajectories to reduce food droppage during transport, which becomes especially critical for foods that can spill or drip (i.e. soup, salsa). We will also extend the system to plan for multiple bites of large food items.

#### ACKNOWLEDGEMENTS

This work was supported by funds from NSF Awards 2132847 and 2218760, as well as the Office of Naval Research.

#### REFERENCES

- [1] S. E. Shune, "An altered eating experience: Attitudes toward feeding assistance among younger and older adults," *Rehabilitation Nursing*, vol. 45, no. 2, pp. 97–105, Apr. 2018. [Online]. Available: <https://doi.org/10.1097/rnj.000000000000147>
- [2] D. Park, Y. K. Kim, Z. M. Erickson, and C. C. Kemp, "Towards assistive feeding with a general-purpose mobile manipulator," *CoRR*, vol. abs/1605.07996, 2016. [Online]. Available: <http://arxiv.org/abs/1605.07996>
- [3] T. Bhattacharjee, E. K. Gordon, R. Scalise, M. E. Cabrera, A. Caspi, M. Cakmak, and S. S. Srinivasa, "Is more autonomy always better? exploring preferences of users with mobility impairments in robot-assisted feeding," in *2020 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2020, pp. 181–190.
- [4] S. Belkhale, E. K. Gordon, Y. Chen, S. Srinivasa, T. Bhattacharjee, and D. Sadigh, "Balancing efficiency and comfort in robot-assisted bite transfer," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 4757–4763.
- [5] <https://meetobi.com/>, [Online; Retrieved on 25th January, 2018].
- [6] 2018, <https://www.secom.co.jp/english/myspoon/food.html>, [Online; Retrieved on 25th January, 2018].
- [7] 2018, <https://www.made2aid.co.uk/productprofile?productid=8&company=RBF%20Healthcare&product=Meal-Mate>.
- [8] 2018, <https://www.performancehealth.com/meal-buddy-system>, [Online; Retrieved on 25th January, 2018].
- [9] M. C. Gemici and A. Saxena, "Learning haptic representation for manipulating deformable food objects," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2014, pp. 638–645.
- [10] R. Feng, Y. Kim, G. Lee, E. K. Gordon, M. Schmitz, S. Kumar, T. Bhattacharjee, and S. S. Srinivasa, "Robot-assisted feeding: Generalizing skewering strategies across food items on a realistic plate," *ISSR*, Jun. 2019.
- [11] E. K. Gordon, X. Meng, T. Bhattacharjee, M. Barnes, and S. S. Srinivasa, "Adaptive robot-assisted feeding: An online learning framework for acquiring previously unseen food items," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2020.
- [12] P. Sundaresan, S. Belkhale, and D. Sadigh, "Learning visuo-haptic skewering strategies for robot-assisted feeding," in *6th Annual Conference on Robot Learning*, 2022. [Online]. Available: <https://openreview.net/forum?id=ILq09gVoaTE>
- [13] J. Grannen, Y. Wu, S. Belkhale, and D. Sadigh, "Learning bimanual scooping policies for food acquisition," in *6th Annual Conference on Robot Learning*, 2022. [Online]. Available: <https://openreview.net/forum?id=qDtbMK67PJG>
- [14] D. Gallenberger, T. Bhattacharjee, Y. Kim, and S. Srinivasa, "Transfer depends on acquisition: Analyzing manipulation strategies for robotic feeding," *ACM/IEEE International Conference on Human-Robot Interaction*, 2019.
- [15] D. Park, Y. K. Kim, Z. M. Erickson, and C. C. Kemp, "Towards assistive feeding with a General-Purpose mobile manipulator," May 2016.
- [16] L. V. Herlant, "Algorithms, Implementation, and Studies on Eating with a Shared Control Robot Arm," Ph.D. dissertation, 2016.
- [17] S. Haddadin and E. Croft, "Physical human–robot interaction," in *Springer handbook of robotics*. Springer, 2016, pp. 1835–1874.
- [18] A. De Santis, B. Siciliano, A. De Luca, and A. Bicchi, "An atlas of physical human–robot interaction," *Mechanism and Machine Theory*, vol. 43, no. 3, pp. 253–270, 2008.
- [19] S. Ikemoto, H. B. Amor, T. Minato, B. Jung, and H. Ishiguro, "Physical human-robot interaction: Mutual learning and adaptation," *IEEE robotics & automation magazine*, vol. 19, no. 4, pp. 24–35, 2012.
- [20] S. Haddadin, A. Albu-Schaffer, A. De Luca, and G. Hirzinger, "Collision detection and reaction: A contribution to safe physical human-robot interaction," in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2008, pp. 3356–3363.
- [21] V. Okunev, T. Nierhoff, and S. Hirche, "Human-preference-based control design: Adaptive robot admittance control for physical human-robot interaction," in *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 2012, pp. 443–448.
- [22] M. Li, A. Canberk, D. P. Losey, and D. Sadigh, "Learning human objectives from sequences of physical corrections," in *International Conference on Robotics and Automation (ICRA)*, May 2021.
- [23] A. Bajcsy, D. P. Losey, M. K. O'Malley, and A. D. Dragan, "Learning robot objectives from physical human interaction," in *Conference on Robot Learning*. PMLR, 2017, pp. 217–226.
- [24] A. Agah and K. Tanie, "Human interaction with a service robot: mobile-manipulator handing over an object to a human," in *Proceedings of International Conference on Robotics and Automation*, vol. 1, April 1997, pp. 575–580 vol.1.
- [25] A. Edsinger and C. C. Kemp, "Human-robot interaction for cooperative manipulation: Handing objects to one another," in *RO-MAN 2007 - The 16th IEEE International Symposium on Robot and Human Interactive Communication*, Aug 2007, pp. 1167–1172.
- [26] M. Huber, M. Rickert, A. Knoll, T. Brandt, and S. Glasauer, "Human-robot interaction in handing-over tasks," in *RO-MAN 2008 - The 17th IEEE International Symposium on Robot and Human Interactive Communication*, Aug 2008, pp. 107–112.
- [27] M. Cakmak, S. S. Srinivasa, M. K. Lee, J. Forlizzi, and

- S. Kiesler, "Human preferences for robot-human hand-over configurations," in *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*. IEEE, 2011, pp. 1986–1993.
- [28] G. Canal, G. Alenyà, and C. Torras, "Personalization framework for adaptive robotic feeding assistance," in *Social Robotics*, A. Agah, J.-J. Cabibihan, A. M. Howard, M. A. Salichs, and H. He, Eds. Cham: Springer International Publishing, 2016, pp. 22–31.
- [29] D. E. King, "Dlib-ml: A machine learning toolkit," *Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009.

## APPENDIX

We provide additional details regarding our approach and experimental evaluation in the sections below. Appendix A highlights the visual differences in transfer trajectory across methods. Appendix B supplements the results from Section IV with a detailed breakdown of user rankings by food type and measured user movement during transfer to further gauge overall, safety, and comfort preferences across methods. Finally, Figure 10 includes visualized rollouts across methods, to further highlight the perceptible differences in transfer. Datasets, code, and videos can be found on our [website](#).

### A. Sample End Effector Trajectories

A difference between the three levels of reactivity evaluated in this paper is clearly seen in the end effector trajectories (see Figure 7). As expected, the less reactive trajectory sticks to its pre-planned path the closest, while the more reactive one jitters away from it. We show trajectories from individual trials because of different user mouth locations and food items between trials. Note that this result is purely qualitative.

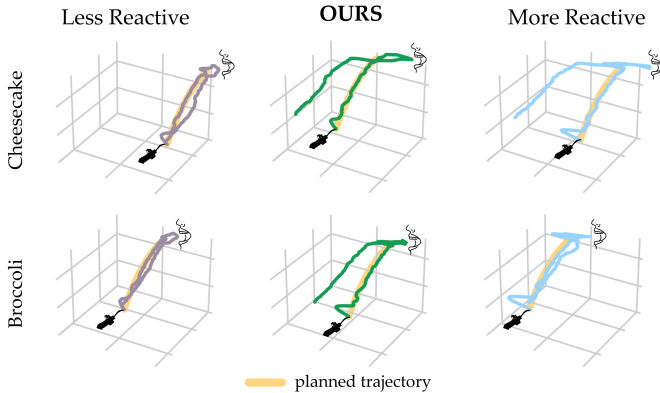


Fig. 7: **End-Effector Trajectories.** Sample end effector position trajectories over the course of a feeding run for the three in-mouth robotic methods. The less reactive trajectory is closest to the pre-planned one, indicating less compliance. In contrast, the more reactive trajectory diverges from the planned one in a jittery path, matching our observations during the trial. Our method balances the two, responding to some user movement without experiencing significant jitter.

### B. Additional User Study Details

In this section, we report additional results from the user study comparing our bite transfer method to three baselines and a human-feeding oracle.

We analyzed our user ranking data, in the range of 1-5 across the safety, comfort, and overall categories, using one-way ANOVA testing to measure the statistical significance of both phased reactivity and transfer style. For phased reactivity, we split the into the categories of *Less Reactive*, *More Reactive*, and *Ours*. For transfer style, we group the *Less Reactive*, *More Reactive*, and *Ours* rankings as the *in-mouth* category, and compare this against the *out-of-mouth Fixed Pose* data. Results for significant findings ( $p\text{-value} \leq 0.05$ ) are reported across the safety, comfort, and overall categories in Tables IV, III, II, respectively. Phased reactivity has a statistically significant effect on all categories of ratings, while

transfer style has a statistically significant affect on comfort ratings alone.

Variable	$p\text{-value}$
Reactivity	<b>0.010802</b>
Transfer-Style	0.949338

TABLE II: **Overall: One-Way ANOVA Findings**

Variable	$p\text{-value}$
Reactivity	<b>0.000011</b>
Transfer-Style	<b>0.019153</b>

TABLE III: **Comfort: One-Way ANOVA Findings**

Variable	$p\text{-value}$
Reactivity	<b>0.000001</b>
Transfer-Style	0.970068

TABLE IV: **Safety: One-Way ANOVA Findings**

In Tables V–IX, we report the values of the 25th, 50th, and 75th percentile of user rankings on safety, comfort and overall bite quality per bite transfer method for each food axis category as shown in Figure 5. As in Figure 5, the 8 food items tested are grouped into three categories for each food variability axis (shape, size, and deformability). These groupings are reported in Table X, which each food item appearing exactly once within each food variability axis classification.

Axis	Category	Safety			Comfort			Overall		
		25th	50th	75th	25th	50th	75th	25th	50th	75th
Shape	Round	2	4	4	2	3	4	2	4	4
	Cube/Cyl	2	4	4	2	2.5	4	2.75	3.5	4
	Irregular	2.75	3.5	4.25	1.75	2.5	3.5	2.75	3	5
Size	Small	2	4	4	2	3	4	2	4	4
	Medium	2	4	4	2	2	4	3	3	4
	Large	2.5	4	4	2	3	4	2.5	3	4.5
Deform	Fragile	1	3	4	1	2	3	1	3	4
	Robust	2	4	4	2	3	4	2	4	5
	Rigid	3	4	4	2	4	4.5	3.5	4	4

TABLE V: **Fixed Pose User Ranking Values.**

Axis	Category	Safety			Comfort			Overall		
		25th	50th	75th	25th	50th	75th	25th	50th	75th
Shape	Round	3	5	5	3	5	5	3	5	5
	Cube/Cyl	3.75	5	5	4	5	5	3.75	5	5
	Irregular	3.75	5	5	4.75	5	5	3.75	5	5
Size	Small	3	5	5	3	5	5	3	5	5
	Medium	5	5	5	4	5	5	5	5	5
	Large	2.5	5	5	3.5	5	5	2.5	5	5
Deform	Fragile	5	5	5	5	5	5	5	5	5
	Robust	3	5	5	4	5	5	3	5	5
	Rigid	2.5	5	5	3	4	5	2.5	5	5

TABLE VI: **More Reactive User Ranking Values.**

Axis	Category	Safety			Comfort			Overall		
		25th	50th	75th	25th	50th	75th	25th	50th	75th
Shape	Round	2	3	4	2	3	4	3	3	4
	Cube/Cyl	2	5	5	3	4	4	2	4	4
	Irregular	2.75	3.5	4	2.75	4	4	2	4	4
Size	Small	2	3	4	2	3	4	3	3	4
	Medium	2	4	4	3	4	4	2	4	4
	Large	2.5	3	4	2.5	4	4	2	4	4
Deform	Fragile	3	4	4	4	4	4	4	4	4
	Robust	2	3	4	2	4	4	2	3	4
	Rigid	2	2	3	2.5	3	3.5	2	2	3.5

TABLE VII: **Less Reactive User Ranking Values.**

Axis	Category	Safety			Comfort			Overall		
		25th	50th	75th	25th	50th	75th	25th	50th	75th
Shape	Round	2	3	3	2	3	3	2	3	3
	Cube/Cyl	2	2.5	3.25	2	3	4	2	2	3
	Irregular	2	2.5	3	3	3	3.25	2	3	3.25
Size	Small	2	3	3	2	3	3	2	3	3
	Medium	2	2	3	2	3	3	2	2	3
	Large	2	3	4	3	3	4.5	2.5	3	4
Deform	Fragile	2	2	3	2	3	3	2	2	2
	Robust	2	3	3	2	3	3	2	3	3
	Rigid	2	4	4.5	2	4	4.5	2	3	4.5

TABLE VIII: OURS User Ranking Values.

Axis	Category	Safety			Comfort			Overall		
		25th	50th	75th	25th	50th	75th	25th	50th	75th
Shape	Round	1	1	1	1	1	1	1	1	1
	Cube/Cyl	1	2	3	1	2	2	1	2	3
	Irregular	1	1	1.25	1	1	1.25	1	1	2
Size	Small	1	1	1	1	1	1	1	1	1
	Medium	1	1	2	1	2	2	1	2	3
	Large	1	2	3	1	1	2	1	1	2
Deform	Fragile	1	2	3	1	2	2	1	2	3
	Robust	1	1	1	1	1	1	1	1	2
	Rigid	1	2	3	1	2	2	1	2	2.5

TABLE IX: Human Oracle User Ranking Values.

Food	Shape	Size	Deformability
Broccoli	Irregular	Large	Robust
Blueberry	Round	Small	Robust
Cherry Tomato	Round	Small	Robust
Strawberry	Irregular	Large	Robust
Carrot	Cylinder	Large	Rigid
Pineapple	Cube	Medium	Rigid
Tofu	Cube	Medium	Fragile
Cheesecake	Cube	Medium	Fragile

TABLE X: Food Properties: We group the 8 food items we consider along three axes of variability: shape, size, and deformability. We consider three classes for each axis, as shown in Figure 5. For shape, we define an “irregularly shaped food” to be one with less than three axes of symmetry.

Figure 9 reports the user rankings for bite transfer methods for each of the 8 foods considered. These results show that our method either outperforms or performs comparably to the More React and Less React baselines in all foods except carrot. Because the carrot is slightly larger than the size of some users mouths, some users preferred more autonomy compared to other foods to adjust the positioning of the carrot and fork in their mouth. As a result, some users preferred the More Reactive baseline over our method as the user could adjust the fork and carrot positioning more easily. Some users also preferred the Less React baseline over our method because the rigidity of the carrot complicates biting off a fork with our method which has a higher reactivity in comparison. As expected, the Human Oracle baseline achieves the highest performance of all the methods across all food items. The Fixed Pose baseline has widely varied performance across the food classes, suggesting some foods, such as tofu cubes and blueberries, are easier to bite off a fixed fork pose, while the other foods are easier to eat when delivered into the user mouth.

We report the user rankings of head movement required for each bite transfer method in Figure 8. We highlight the Fixed Pose method required the most head movement from the user, due to the nature of the fork approaching the mouth position, but stopping outside instead of attempting

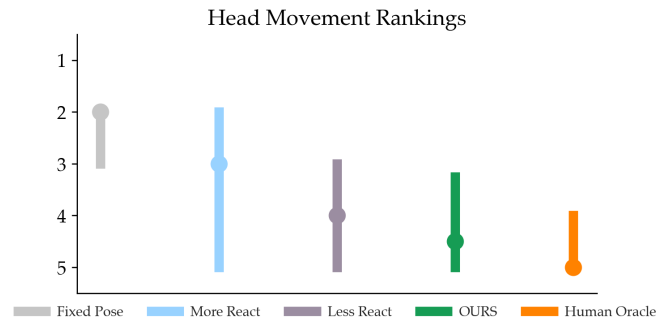


Fig. 8: Head Movement User Rankings. We report the user rankings for the ease with which they could take a bite with each bite transfer method without moving their head. 1 signifies the user had to move more to take a bite, while 5 signifies less head movement. We find that the Fixed Pose baseline required the most head movement due to the nature of the method holding the food item outside the mouth for the user to take a bite.

an in-mouth transfer. The three in-mouth robotic bite transfer methods (More React, Less React, and Ours) are all ranked similarly for required head movement, which is expected as these methods use the same transfer trajectory to achieve in-mouth delivery. Additionally, we note that our bite transfer method approaches the Human Oracle performance in terms of amount of head movement required.

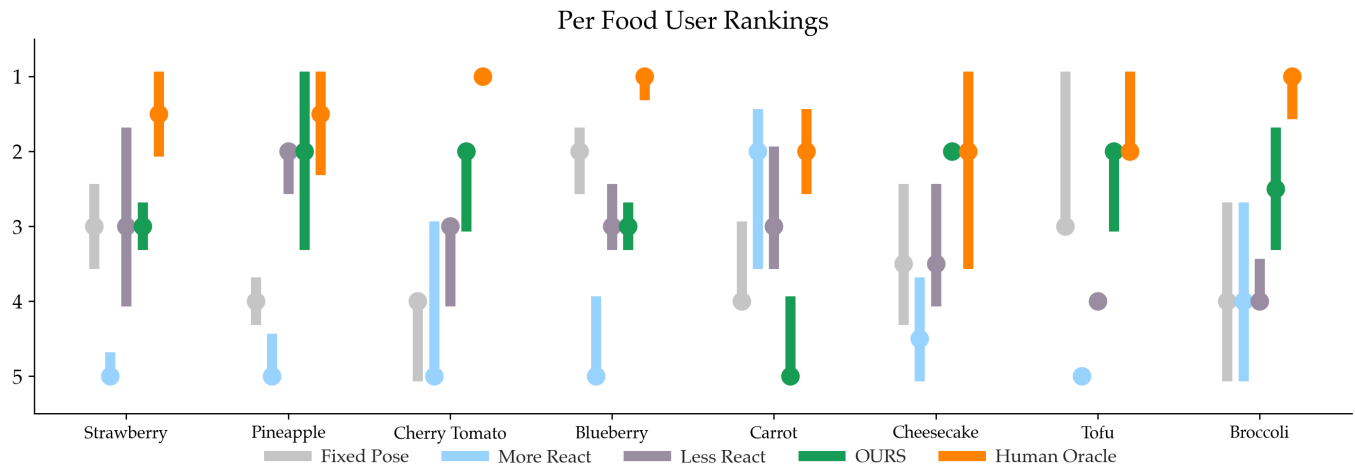


Fig. 9: **Per Food User Rankings.** We report the median and interquartile range of user rankings of the overall bite quality for our bite transfer method and four baselines for each of the 8 food groups considered. 1 indicates best performance and 5 indicates worst. We observe that our method performs as well or better than all robotic feeding baselines across all foods except carrots. We highlight that our method approaches human feeding oracle performance in the pineapple and cheesecake food classes, supporting the effectiveness of our robotic feeding method.

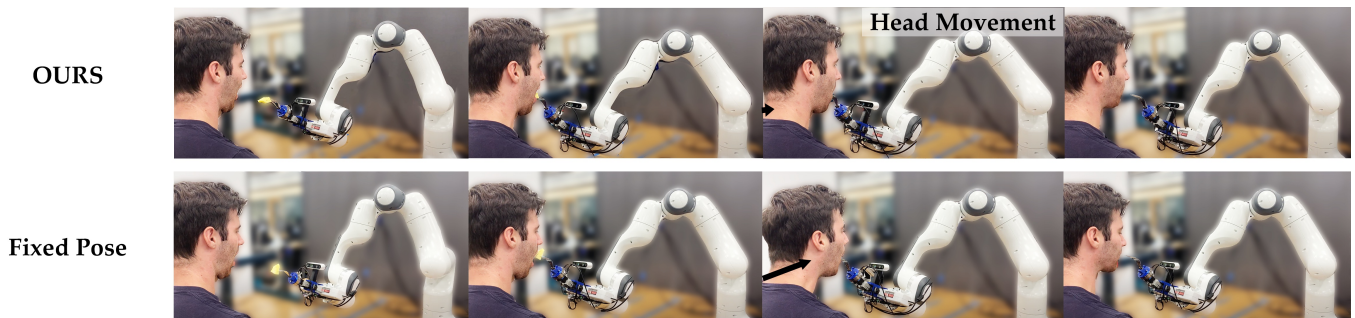


Fig. 10: **Sample Trajectories.** We show two transfer trajectories for pineapple with *Ours* and *Fixed Pose* methods. With reactive, *in-mouth* transfer, *Ours* requires less head movement from the user to successfully transfer the bite compared to the *out-of-mouth Fixed Pose* strategy.