FRANC: Feeding Robot for Adaptive Needs and Personalized Care

J-Anne Yow[1,2,*], Luke Thien Luk Toh[1,*], Yi Heng San[1], Wei Tech Ang[1,2]
[*]These authors contributed equally

***Abstract*—Robot-assisted feeding systems have the potential to significantly enhance the independence and quality of life of individuals with mobility impairments. While prior work has focused on personalizing bite sequences based on user feedback provided only at the start of the feeding process, this approach assumes that users can fully articulate their preferences upfront. In reality, it is cognitively challenging for users to anticipate every detail, and their preferences may evolve during feeding. Thus, there is a need for an adaptive system that supports iterative corrections across all stages of the feeding process while maintaining context and feeding history to interpret inputs relative to earlier instructions. In this paper, we present FRANC, a novel framework for personalized RAF that leverages large language models (LLMs) with a decomposed prompting strategy to dynamically adjust bite sequence, acquisition and transfer parameters during feeding. Our approach allows iterative corrections without sacrificing consistency and accuracy. In our user studies, FRANC improved bite sequencing accuracy from 65% to 93% and enhanced user satisfaction, with participants reliably perceiving when their preferences were being integrated despite occasional execution failures. We also provide a detailed failure analysis and offer insights for developing more adaptive and effective robot-assisted feeding systems.**

***Index Terms*—Robot-Assisted Feeding, Personalization, Large-Language Models, Assistive Robotics**

I. Introduction

Robot-assisted feeding (RAF) systems have the potential to increase independence and improve the quality of life of people with mobility impairments while reducing caregiver burden. Feeding is an inherently personal activity, requiring adaptation to individual needs and preferences. Just as human caregivers tailor feeding to each person, RAF systems must recognize and respond to user preferences dynamically. However, achieving such personalization remains a challenging problem in robotics.

Personalization in RAF systems is particularly complex as it necessitates adaptation across all stages of the feeding process — *bite selection*, where the robot decides which piece of food to pick up; *bite acquisition*, where the robot picks up the food; and *bite transfer*, where the robot delivers the food to the user's mouth. Recent works have addressed specific aspects of the feeding process, such as determining effective skewering strategies during food acquisition [1]–[5], and developing reliable bite transfer strategies [6]–[11], but they generally do not allow users to specify and modify their preferences. More recently, FLAIR [12] integrated user preferences for bite sequence planning, allowing users to provide language feedback at the beginning of the feeding sequence. However, this approach imposes a significant cognitive burden, especially since users may not fully know their preferences in advance, and prevents adaptive corrections during feeding.

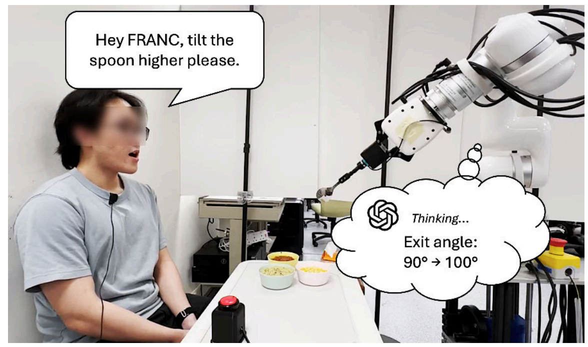

Fig. 1: A user provides language feedback and FRANC adjusts its feeding parameters to match the user's preference.

To the best of our knowledge, no prior work has proposed a comprehensive system that integrates user preferences across all stages of feeding while supporting multiple corrections — a gap largely driven by the complexity of managing such an integrated process. Our work assumes that user preferences are provided via natural language inputs and leverages large language models (LLMs) for planning. We chose natural language as the interaction mode because it offers a natural and intuitive way for users to express complex, nuanced preferences. However, this flexibility comes at the cost of increased ambiguity and contextual dependence. As the task complexity increases with the inclusion of more feeding stages, so does the complexity of the prompts required to accurately convey user intent, which can reduce the accuracy of preference interpretation and task execution.

Furthermore, allowing users to provide multiple corrections or interruptions mid-process introduces its own challenges. The system must maintain context and track feeding history to

The research was conducted at the Future Health Technologies at the Singapore-ETH Centre, which was established collaboratively between ETH Zurich and the National Research Foundation Singapore. This research is supported by the National Research Foundation Singapore (NRF) under its Campus for Research Excellence and Technological Enterprise (CREATE) programme.

[1]School of Mechanical and Aerospace Engineering, Nanyang Technological University Singapore
[2]Singapore-ETH Centre, Future Health Technologies Programme
Corresponding author's email: janne.yow@ntu.edu.sg

ensure that subsequent actions are both correct and consistent with previous inputs. Although LLMs have demonstrated impressive natural language understanding and reasoning capabilities, their effectiveness is highly dependent on prompt structure and design.

In this paper, we introduce **FRANC – Feeding Robot for Adaptive Needs and Personalized Care**, a RAF system that personalizes the entire feeding process – from bite sequencing to acquisition and transfer – using natural language feedback. To ensure safe and predictable robot behaviour in assistive settings where close physical interaction is required, we intentionally constrain the space of personalization to a predefined set of interpretable parameters. FRANC leverages LLMs to interpret language feedback and generate parameters for its library of feeding primitives, enabling the robot to dynamically adjust its behavior during feeding while supporting iterative corrections. We employ a decomposed prompting approach that breaks down the overall task into manageable components. This strategy improves the accuracy of language-based interpretation and enhances the consistency of the robot's actions, even when users provide multiple corrections throughout the feeding cycle.

From the insights gained in preliminary experiments, we refined our approach and increased bite sequencing accuracy from 65% to 92.86%. User studies with a real robot also validate FRANC's effectiveness, demonstrating that users can reliably detect when the robot adapts to their preferences and that such personalization enhances overall satisfaction.

Our key contributions include:

- **FRANC – A Personalized Feeding System:** We present an RAF system that personalizes the entire feeding process, spanning bite sequencing, acquisition, and transfer, using natural language feedback. FRANC supports iterative, adaptive corrections throughout the feeding cycle, enabling users to refine their preferences in real-time.
- **User Studies:** We conduct a user study to evaluate the impact of language-driven personalization on user satisfaction, system responsiveness, and overall performance.
- **Failure Analysis and Insights:** We provide a detailed analysis of failure cases and lessons learned from our user studies, offering insights for further improvements in robot feeding assistants.

II. Related Work

A. Robot-Assisted Feeding

The robot-assisted feeding (RAF) process consists of three stages – bite selection, bite acquisition and bite transfer. The feeding sequence typically begins when the user signals the system, with some systems also allowing users to select the food item to be eaten. While many studies assume automatic bite selection [1]–[6], [13], [14], often aimed at clearing the plate efficiently, user studies have highlighted the importance of user control over bite selection and initiation [15], [16]. Thus, recent works have explored adaptive bite sequencing strategies to better align with user preferences [12].

Bite acquisition has primarily focused on using forks for skewering various food items [1]–[6], [13]. The robot must determine a suitable skewering strategy (e.g. vertical or angled) to acquire food successfully. While vision-based approaches have been effective for foods with distinguishable shapes [2], [4], [6], they struggle to generalize to foods with similar appearances but varying textures, such as tofu of different hardness. To address this limitation, haptic feedback has also been used to determine the hardness of the food and guide skewering strategies [1], [3], [5], [13].

The use of spoons to acquire foods such as yogurt and cereal has also been studied, but rely on a predefined manipulation strategy [17]. Vision-based classifiers have been trained to assess whether an adequate amount of food was scooped [18], but these approaches do not adjust the amount of food acquired. To address challenges with fragile foods (e.g., tofu, cheesecake) or rolling foods (e.g., peas, macaroni), bimanual scooping techniques have been proposed to stabilize items during acquisition [19]. However, varying the amount of food scooped remains underexplored, despite user feedback suggesting that this could significantly improve satisfaction [16], [17].

In bite transfer, early works focused on optimizing the angle of approach to maximize user comfort and positioning food near the mouth [6]. For individuals with severe mobility limitations, in-mouth bite transfer has been addressed using compliant controllers based on force feedback to enhance comfort and safety [9], [20]. Furthermore, the bite transfer trajectory has been parameterized into distinct components, including entry angle, entry depth, exit angle, and exit depth, to study how these parameters affect user comfort [10], [11].

Most prior works address specific parts of the feeding process, focusing on individual stages such as bite selection, acquisition, or transfer, without incorporating user feedback to modify their preferences. Some studies present comprehensive systems that span the entire feeding process [17], [21], [22], but they also do not support personalization. A fully personalized RAF system that adapts to user preferences across all stages, while also allowing for iterative corrections during the feeding process remains underexplored.

B. Adapting to User Preferences

Adapting to user preferences is crucial for personalized robot-assisted systems. Early works provide general taxonomies for assistive robots [23], while more domain-specific efforts focus on taxonomies for bite acquisition strategies using a fork [1]. Frameworks for adapting task and motion plans have been explored in other domains, such as shoe-dressing [24], where systems personalize both action sequences and motion trajectories to align with user needs.

In the domain of feeding, personalization frameworks have enabled adaptive robotic trajectories from caregiver demonstrations [23], while recent studies highlight the importance of balancing autonomy and user control [25]. More recently, the emergence of large language models (LLMs) has introduced new capabilities for interpreting and responding to user preferences. For example, ExTraCT has demonstrated adaptive bite size and feeding speed based on language input [26], but it relied on predefined templates to map user preferences to robot actions. Similarly, FLAIR [12] introduced adaptive bite sequencing using language feedback; however, its scope was limited to bite sequencing and did not accommodate iterative corrections. This approach assumes that users can plan all their preferences in advance, but our user study reveals that users typically consider only the immediate next few bites, as planning more steps ahead can be cognitively challenging. Thus, there is a need for a RAF system that supports multiple, adaptive corrections during the feeding cycle.

III. Problem Formulation

To allow a robot to adapt its assistance based on language feedback, we assume access to a library of skills the robot can execute to perform the feeding process. Each skill s represents a motion primitive parameterized by a set of values p, which define the trajectory to be executed. For example, a `scoop` skill takes inputs such as the food pose and bite-size and outputs the trajectory required to scoop the food item of choice with the desired bite size. Given an optional language feedback, our goal is to determine the appropriate parameters for each skill primitive to fulfill the user's preferences.

A. Skill Library

To demonstrate personalization across the feeding pipeline, we modify preferences across the bite selection, bite acquisition, and bite transfer stages.

For bite selection and acquisition, we use a goal-conditioned `scoop` primitive learned using Reinforcement Learning (RL) [27]. This primitive is conditioned on a target amount and is parameterized by:

- **Food pose** (x_f, y_f, z_f): The position of the food to be scooped, first determined by the bite sequence. Given an RGB image of the food and the corresponding labels, we use Grounding DINO [28] to get the corresponding bounding box and food pose.
- **Bite size** (a): The amount of food to be scooped, controlled by setting the target amount for the goal-conditioned scooping policy. (range: 1–5, with 1 being the smallest bite size; default: 3)

For bite transfer, we define a `transfer` primitive to deliver the food to the user, similar to [10]. This is parameterized by:

- **Distance to mouth** (d): The distance at which the robot delivers the food relative to the user's mouth. (range: 5–10 cm; default: 7.5 cm)
- **Exit angle** (α): The tilt applied to the spoon after a bite has been taken to ensure comfort. (range: 80°–110°; default: 90°)
- **Transfer speed** (v): The speed at which the arm moves when delivering food. (range: 1–10, default: 5)

TABLE I: User preference categories with corresponding frequency and accuracy

Categories	Frequency	Accuracy
Alternating between all food items	17%	52.94%
Alternating X with other food items	24%	45.83%
Finish one food item before next	13%	84.62%
Avoid food item	6%	100.00%
No preference	4%	75.00%
Combination preference	36%	69.44%

B. Mapping User Preferences to Motion Parameters

Building on the skill library, the next objective is to translate user-provided natural language inputs into concrete parameter values for each motion primitive. We first adopted FLAIR [12], modifying it to incorporate the parameters defined above (FLAIR-modified). To assess the feasibility of this approach and gather insights, we conducted a preliminary user study.

IV. Preliminary User Study

While motion primitives can often be evaluated objectively by measuring parameter accuracy, evaluating bite sequencing is more challenging due to the inherent variability in user preferences and language ambiguity. To address these challenges, the preliminary study had two primary objectives: (1) to evaluate the baseline performance of FLAIR-modified in interpreting natural language preferences for bite sequencing, and (2) to collect language inputs and corresponding bite sequence pairs that reveal these ambiguities.

Conducting an offline evaluation before real-robot deployment allowed us to assess how well the system could translate user language inputs into bite sequences without the added complexity of real-time execution. The Nanyang Technological University Institutional Review Board (IRB-2023-571) approval was obtained for this study.

A. Study Design

A total of 10 participants (10 male, age range 22–30) were recruited. They were shown images of ten plates of food and asked to specify their preferred bite sequence in natural language. The system generated a corresponding sequence for each plate, and participants rated how closely it matched their expectations using a four-point Likert scale — "Exactly as expected", "Mostly as expected", "Somewhat different from expected", or "Completely different from expected". A neutral option was excluded to ensure participants made a definitive evaluation of each sequence.

B. Results and Insights

We defined an accurate output as one that received a rating of "Exactly as expected" or "Mostly as expected". Overall, only 65% of the generated bite sequences were deemed accurate. To gain deeper insights into user preferences and common failure modes, we categorized the language inputs (Table I).

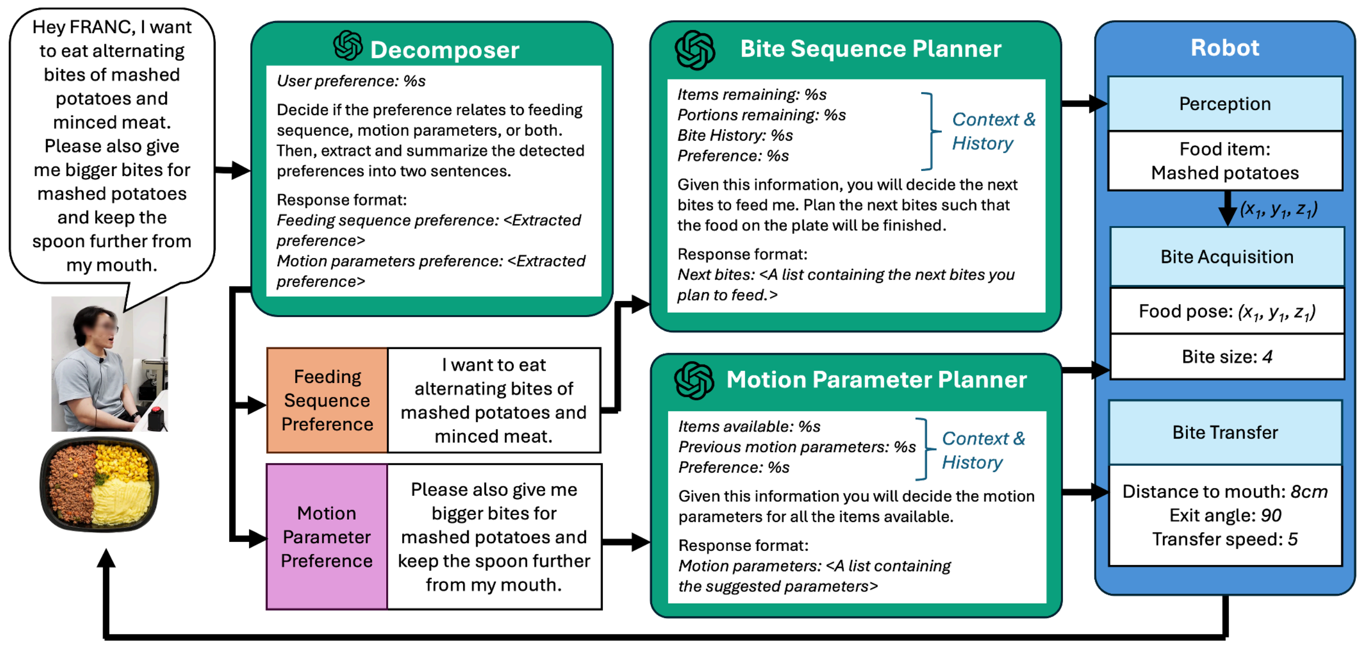

Fig. 2: Overview of FRANC. The system first decomposes the user's language input into components for bite sequencing and motion parameter adjustments. The bite sequence planner then generates the complete sequence, while the motion parameter planner produces the appropriate, food-specific parameters for bite acquisition and transfer based on the library of skills available to FRANC. (Prompts are abbreviated for clarity.)

Notably, the system consistently succeeded in avoiding a particular food, but struggled the most with generating alternating bite sequences, especially in the categories "Alternating one food item with others" (e.g. *"Alternate rice and chicken"*) and "Alternating X with other food items" (e.g. *"Alternate rice with the other food items"*). These insights guided our subsequent system improvements, as described in Section V.

We observed several language ambiguities that led to distinctly different desired bite sequences. For example, one participant who preferred to finish each food item completely before moving on stated: "Vegetables first, then rice, then egg, then fish, please." *(Desired sequence: ['green beans', 'green beans', 'green beans', 'rice', 'rice', 'rice', 'egg', 'egg', 'egg', 'fish', 'fish', 'fish']).* In contrast, another participant with a preference for alternating between food items provided: "Carbs first, protein second, eggs before fish, no beans." *(Desired sequence: ['rice', 'egg', 'fish', 'rice', 'egg', 'fish', 'rice', 'egg', 'fish']).* Although both statements share a similar sentence structure, they reflect markedly different sequencing intentions, highlighting how similar language can imply different meanings.

V. Approach

Building on the insights from our preliminary user study, we refined our method to address observed failure cases and better capture user intent. In this section, we introduce **FRANC — Feeding Robot for Adaptive Needs and Personalized Care**, a comprehensive RAF system that employs a decomposed prompting strategy to enhance the consistency and accuracy of both bite sequence generation and motion parameter outputs. Our approach supports iterative corrections throughout the feeding cycle, enabling users to modify their preferences dynamically as they interact with the robot.

A. Personalization via Foundation Models

To adapt to user preferences throughout the feeding process, we leverage the reasoning capabilities of large language models (LLMs), specifically GPT-4o. The LLM interprets natural language feedback in the context of meal details and feeding history, generating appropriate parameters for each skill primitive detailed in Section III-A.

Due to the complexity required in language understanding when including bite acquisition and bite transfer parameters into a single prompt, the performance of FLAIR-modified degraded. To overcome this challenge, we adopted a decomposed prompting strategy [29] that breaks the overall task into modular queries, each targeting a specific stage.

Our system, **FRANC** (Fig. 2), employs a two-step process:

1) Input Decomposition: Given a language input, a decomposer classifies it into two categories: bite sequencing and motion parameter adjustment. Additionally, the decomposer segments the input into relevant components, which are then routed to the appropriate planner.

2) Specialized Planners:

- A **Bite Sequence Planner** generates the complete bite sequence in a single inference — unlike FLAIR, which relies on iterative prompting after each bite — resulting in more consistent and coherent sequencing outputs.

FLAIR-modified (No decomposer)

	Bite Sequence	Bite Size	Distance to mouth	Exit Angle	Transfer Speed
1	broccoli	**1.0**	5.0	80.0	5.0
2	rice	3.0	5.0	80.0	5.0
3	broccoli	3.0	5.0	80.0	5.0
4	rice	3.0	5.0	80.0	5.0
5	beef	3.0	5.0	80.0	5.0
6	rice	3.0	*7.5*	80.0	5.0
7	broccoli	3.0	*7.5*	*90.0*	5.0
8	**orange**	3.0	*7.5*	*90.0*	5.0
9	**beef**	2.0	8.0	*90.0*	5.0
10	**beef**	*1.0*	*8.5*	*90.0*	5.0
11	orange	2.0	*9.0*	*90.0*	5.0
12	orange	*1.0*	*9.0*	*90.0*	5.0

Food items:
rice
beef
broccoli
oranges

Preference #1

Hey FRANC, I want to alternate between rice and the other foods. Please keep the spoon closer to my mouth and feed me with a low tilt in the spoon.

Preference #2

Hey FRANC, I would instead like to have oranges as desert and alternate between what's remaining.

Preference #3

Hey FRANC, The spoon is too close, can you keep the spoon further from my mouth and I'm getting full, so please feed me smaller bites now.

FRANC (Decomposer)

	Bite Sequence	Bite Size	Distance to mouth	Exit Angle	Transfer Speed
1	broccoli	3.0	7.0	85.0	5.0
2	rice	3.0	7.0	85.0	5.0
3	broccoli	3.0	7.0	85.0	5.0
4	rice	3.0	7.0	85.0	5.0
5	beef	3.0	7.0	85.0	5.0
6	rice	3.0	7.0	85.0	5.0
7	beef	3.0	7.0	85.0	5.0
8	broccoli	3.0	7.0	85.0	5.0
9	beef	1.0	8.0	85.0	5.0
10	orange	1.0	8.0	85.0	5.0
11	orange	1.0	8.0	85.0	5.0
12	orange	1.0	8.0	85.0	5.0

Fig. 3: A comparison of FLAIR-modifed vs FRANC in suggesting parameters based on multiple corrections during the same feeding session. The bolded values in red indicate an incorrect parameter, while the italicized values in yellow indicate inconsistent parameters. The default values are in Section III-A.

- A **Motion Parameter Planner** produces the parameters for bite acquisition and transfer, including bite size, transfer distance, exit angle, and transfer speed. The LLM is also prompted to articulate its reasoning, ensuring that the suggestions are context-aware. A distinct set of parameters is generated for each food item, enabling the system to tailor its motion strategies to the unique properties and user preferences associated with each type of food.

For improved performance, all prompts are augmented with few-shot examples [30], [31] that illustrate the desired output format and reasoning process. To ensure consistency while allowing for iterative language corrections, each planner is provided with the relevant context and feeding history. The full prompts are available on our website.

When a new language correction is received during feeding, the system selectively updates only the affected components, invoking the bite sequence planner for sequence modifications and the motion parameter planner for adjustments to acquisition and transfer. This selective updating not only enhances computational efficiency but also maintains consistency with the feeding history, enabling robust, personalized adaptation throughout the entire feeding cycle. The output difference between FLAIR-modified and FRANC is shown in Fig. 3.

VI. Experiments

In this section, we evaluate FRANC's ability to adapt to user preferences based on language inputs. Our experiments aim to address two key questions:

- **RQ1:** Can users reliably identify when the robot adapts to their language inputs?
- **RQ2:** Does personalization via language feedback enhance the overall user experience?

A. User Study Setup and Design

We recruited 7 participants without mobility limitations (7 male, age range 22–43), after obtaining approval from the Nanyang Technological University Institutional Review Board (IRB-2023-571). The experimental setup comprised an xArm-6 robotic arm (UFactory, China) equipped with an IKEA Fornuft spoon attached to a utensil mount [32] and a 6-axis force-torque sensor (M3712B, Sunrise Instruments, China). A Realsense D435i RGB-D camera (Intel, USA) was mounted on the gripper to facilitate perception. The robot was positioned in front of a table with three bowls containing mashed potatoes, corn, and minced meat (see Fig. 1).

During the feeding process, the robot provided verbal instructions and feedback via a text-to-speech module, informing participants when to provide their preferences, indicating which food was being acquired, and signaling when feeding and eating would occur. Due to challenges with speech recognition for local English accents, participants' language feedback was manually entered into the system by the experimenter.

To manage the complexity and time required for full-system evaluations, we split the study into two distinct phases.

1) Phase 1 – Bite Sequence Evaluation: Participants provided their feeding preferences and preferred bite sequences for four different plates of food. These language inputs were then fed into FRANC to generate a corresponding feeding sequence. If the generated sequence did not exactly match their expectations, participants were asked to provide revised language input along with their expected sequence. This phase enabled us to evaluate bite sequence preferences across varied plates without the additional time complexity of testing the full feeding system.

2) Phase 2 – Full System Evaluation: Participants were seated in front of the robotic arm to evaluate the complete feeding system. After a practice session of three feeding cycles to familiarize themselves with the system, participants completed nine feeding cycles under two scenarios (18 cycles total):

- **No-Adaptation**: This baseline represents the current state of the art in RAF, where there is no language-based adaptation for bite acquisition and transfer. In our study, a predetermined bite sequence was used.
- **FRANC**: Our proposed approach that adapts to user preferences across the whole feeding pipeline.

The order of scenarios was counterbalanced to reduce order effects. At the beginning of each feeding cycle, participants provided their feeding preferences and had the option to provide more language inputs after each bite. After each scenario, they completed an evaluation form to assess the system's performance. Participants rated their agreement on the following statements on a 5-point Likert scale: (1) I was satisfied with the overall feeding experience (satisfaction). (2) The robot was following my preferences accurately (preferences). (3) The behavior of the robot was significantly different when I gave a preference (responsiveness). (4) I felt comfortable during the feeding process in this scenario (comfort). (5) I felt safe during the feeding process in this scenario (safety). (6) It was easy to communicate my preferences and interact with the robot during the feeding process (ease of use).

At the conclusion of the trials, participants indicated which scenario they believed demonstrated the robot's adaptation to their preferences, selected their overall preferred scenario, and ranked which aspects of the feeding process were most critical for adaptation. Additionally, they provided qualitative feedback on other preferences they would like the robot to adapt to, as well as any general comments.

VII. Results

Due to a malfunction in the robot's force-torque sensor during Phase 2 of one participant's session, we omitted that participant's Phase 2 data from our analysis.

In Phase 1 of this experiment, FRANC generated bite sequences with 92.86% accuracy, a significant improvement over the 65% accuracy observed in the preliminary study using FLAIR-modified. This result confirms that our approach enhances the translation of natural language inputs into accurate bite sequences.

In Phase 2, FRANC outperformed No-Adaptation across several key metrics (Fig. 4), particularly in **responsiveness**, **satisfaction**, and adherence to their **preferences**. When asked to indicate their overall preference, all six participants (with complete data) chose FRANC over the No-Adaptation baseline. Even though the robot did not always execute actions perfectly due to occasional failure cases and inaccuracies, five out of the six participants still accurately identified that the robot was following their preferences in the FRANC scenario. This indicates that the system's adaptive behavior was sufficiently noticeable for users to distinguish it from a non-adaptive baseline, despite imperfections in the system (**RQ1**).

Participants provided an average of 6 language inputs for both No-Adaptation and FRANC, highlighting the importance of a system that supports iterative corrections. Furthermore, participants agreed that a system capable of adapting to their preferences would enhance the overall feeding experience, with an average rating of 4.67 out of 5 on this metric. These results suggest that our personalized, adaptive approach not only improves the accuracy of bite sequencing and motion parameter estimation but also leads to higher user satisfaction and perceived system responsiveness (**RQ2**).

However, participants rated No-Adaptation slightly higher in terms of **safety** compared to FRANC. This difference was primarily attributed to larger, more abrupt movements during the bite transfer phase when users requested higher exit angles. For instance, if a participant requested a higher exit angle while seated at a distance, one of the robot's joints would execute a large motion, which although not causing any injuries, was perceived as less safe. Despite these instances, participants generally found the feeding process **safe**, **easy to use** and **comfortable**.

Participants were also asked to rank the importance of robot adaptation for various phases of the feeding process — bite sequence, bite size, bite transfer parameters, and others. Interestingly, no single parameter dominated the rankings, suggesting that participants found multiple aspects of feeding important to personalize. This diversity in preferences underscores the need for a system that can adapt to various facets of the feeding process.

A. Failure Analysis

Across the user study, we recorded 32 failure cases—some resulting in a failed feeding cycle, while others were errors that did not prevent the cycle from completing. We classified them into four categories: *perception* failures, which arise from inaccuracies in food bowl detection and localization; *language understanding* failures, where there is a mismatch between user intent and the LLM-generated bite sequence and motion parameters; *bite acquisition* failures, where the robot is unable to effectively scoop the food from the bowl; and *bite transfer* failures, which occur when the robot cannot accurately deliver food to the user's mouth. Table II shows the relative failure rates for each category.

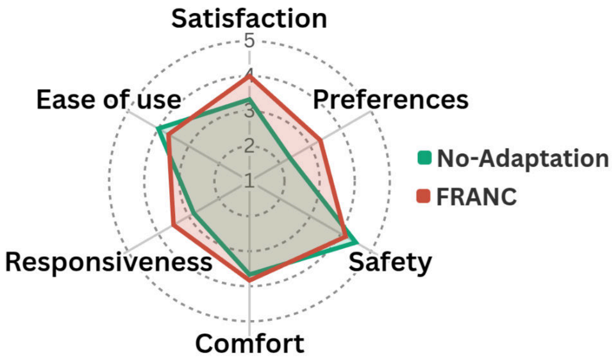

Fig. 4: Subjective ratings from the user study

TABLE II: Failure analysis: relative failure rates by category.

Categories	Relative Failure Rate
Perception	9.4%
Language Understanding	18.8%
Bite Acquisition	40.6%
Bite Transfer	31.3%

Bite acquisition and *bite transfer* collectively account for over two-thirds of all failures. Analysis indicates that these issues largely stem from the robot's kinematic limitations and workspace constraints. For instance, when the food bowl is positioned near the edge of the robot's workspace or the user's position is outside the optimal range (e.g., when the user leans backwards), the robot struggles to achieve the necessary joint configurations for effective scooping and transfer, resulting in abrupt, large motions or failure in execution. Furthermore, for smaller bite sizes, the goal-conditioned scooping policy sometimes fails to acquire food. Some errors were also due to a failure to detect the spoon, required for the scooping primitive. For bite transfer, some failures were due to the system's inability to detect when a user had already taken a bite, particularly in cases of gentle biting.

Language understanding failures primarily occurred due to ambiguities in user instructions, resulting in misclassification by the decomposer. For example, when a subject stated "Give me bigger bites of minced meat", the system interpreted it solely as an instruction to increase the bite size for minced meat, even though, given the context, a human would likely interpret the command as both increasing the bite size and ensuring that minced meat is fed next.

B. Lessons Learned and Future Improvements

To reduce the failure rates and enhance the overall feeding experience, we suggest several areas of improvement:

1) Workspace optimization: Optimize the feeding setup to ensure that the food bowls and user are consistently positioned within the robot's optimal workspace. This could involve marking a defined workspace area on the table for the bowls and providing visual or auditory cues if users go beyond the workspace. This would also help mitigate issues related to joint configuration constraints and reduce abrupt movements during bite transfer.

2) Safety checks: Incorporate additional mechanisms to determine when the user's mouth is within the optimal range, so that real-time feedback can alert users if they stray too far, ensuring safe and effective bite transfer. Moreover, integrating human state understanding could improve bite transfer timing by detecting when a user is about to take a bite and when a bite has already been taken, especially in cases of gentle biting, which our current system sometimes fails to detect.

3) Interactive verbal communication: Incorporating interactive verbal communication so that FRANC can ask clarifying questions when user commands are ambiguous, and allowing users to set their desired verbosity to adjust the frequency and type of verbal feedback provided by the robot.

4) Improved user interface: Supplement verbal feedback with additional indicators (e.g., colored lights or display messages on a graphical user interface) to communicate the robot's status and decision-making process. This multimodal feedback would help users better understand system behavior and know when to prepare for subsequent actions.

5) Bite-size consistency: Address variations in bite acquisition across different food types. For example, an increased bite size request for mashed potatoes may work well, but applying the same adjustment to minced meat can yield excessively large bites. Fine-tuning the goal-conditioned scooping policy to incorporate food-specific characteristics is essential for achieving consistent bite sizes.

6) Speech-to-text understanding for local language: We attempted to use an off-the-shelf speech-to-text module to parse user instructions; however, it struggled with local English accents. Future work will explore modules trained with local accents or adapt existing models to better accommodate these linguistic variations, ensuring more robust language feedback.

VIII. Conclusion

We presented FRANC — Feeding Robot for Adaptive Needs and Personalized Care — a robot-assisted feeding system that personalizes the entire feeding process through natural language feedback. Our work demonstrates that by coupling natural language feedback with a decomposed prompting strategy, FRANC can effectively adapt its behavior across all stages of the feeding process. FRANC dynamically adjusts bite sequencing, acquisition, and transfer parameters during the feeding process, supporting iterative corrections as users refine their preferences. Our experimental results demonstrate significant improvements in bite sequencing accuracy (from 65% to 92.86%), enhanced user satisfaction, and a robust ability for users to perceive when their preferences are being integrated.

While these outcomes are promising, our study also revealed several insights that highlight opportunities for future enhancement. Safety concerns emerged due to abrupt and large joint motions during bite transfer, and occasional execution inaccuracies indicate that our current control and perception modules require further refinement. Additionally, our user studies highlight the need for improved human-robot interaction, particularly through more interactive communication, and enhanced bite size consistency across different food types.

In summary, although further improvements are needed, FRANC represents a significant step toward creating a more personalized and adaptive robot-assisted feeding system. Future work will focus on enhancing safety and robustness, followed by more in-depth user studies involving individuals who require feeding assistance, to validate and extend FRANC's capabilities in real-world assistive contexts.

References

[1] T. Bhattacharjee, G. Lee, H. Song, and S. S. Srinivasa, "Towards robotic feeding: Role of haptics in fork-based food manipulation," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1485–1492, 2019.

[2] 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," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 9659–9666.

[3] E. K. Gordon, S. Roychowdhury, T. Bhattacharjee, K. Jamieson, and S. S. Srinivasa, "Leveraging post hoc context for faster learning in bandit settings with applications in robot-assisted feeding," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 10 528–10 535.

[4] R. Feng, Y. Kim, G. Lee, E. K. Gordon, M. Schmittle, S. Kumar, T. Bhattacharjee, and S. S. Srinivasa, "Robot-assisted feeding: Generalizing skewering strategies across food items on a plate," in *Robotics Research: The 19th International Symposium ISRR*. Springer, 2022, pp. 427–442.

[5] P. Sundaresan, S. Belkhale, and D. Sadigh, "Learning visuo-haptic skewering strategies for robot-assisted feeding," in *6th Annual Conference on Robot Learning*, 2022.

[6] D. Gallenberger, T. Bhattacharjee, Y. Kim, and S. S. Srinivasa, "Transfer depends on acquisition: Analyzing manipulation strategies for robotic feeding," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2019, pp. 267–276.

[7] 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 *Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction*, 2020, pp. 181–190.

[8] 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.

[9] L. Shaikewitz, Y. Wu, S. Belkhale, J. Grannen, P. Sundaresan, and D. Sadigh, "In-mouth robotic bite transfer with visual and haptic sensing," in *2023 International Conference on Robotics and Automation (ICRA)*. IEEE, 2023.

[10] Y. H. San, V. Ravichandram, J.-A. Yow, S. S. Chan, Y. Wang, and W. T. Ang, "Simulating safe bite transfer in robot-assisted feeding with a soft head and articulated jaw," in *2025 International Conference On Rehabilitation Robotics (ICORR)*. IEEE, 2025, pp. 628–633.

[11] S. S. Chan, J.-A. Yow, Y. H. San, V. Ravichandram, Y. Wang, L. S. Lim, and W. T. Ang, "Orbit: Optimizing robot-assisted bite transfer leveraging a real2sim2real framework," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2025.

[12] R. K. Jenamani, P. Sundaresan, M. Sakr, T. Bhattacharjee, and D. Sadigh, "Flair: Feeding via long-horizon acquisition of realistic dishes," *arXiv preprint arXiv:2407.07561*, 2024.

[13] H. Song, T. Bhattacharjee, and S. S. Srinivasa, "Sensing shear forces during food manipulation: resolving the trade-off between range and sensitivity," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 8367–8373.

[14] P. Sundaresan, J. Wu, and D. Sadigh, "Learning sequential acquisition policies for robot-assisted feeding," *arXiv preprint arXiv:2309.05197*, 2023.

[15] T. Bhattacharjee, M. E. Cabrera, A. Caspi, M. Cakmak, and S. S. Srinivasa, "A community-centered design framework for robot-assisted feeding systems," in *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility*, 2019, pp. 482–494.

[16] A. Nanavati, P. Alves-Oliveira, T. Schrenk, E. K. Gordon, M. Cakmak, and S. S. Srinivasa, "Design principles for robot-assisted feeding in social contexts," in *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, 2023, pp. 24–33.

[17] D. Park, Y. Hoshi, H. P. Mahajan, H. K. Kim, Z. Erickson, W. A. Rogers, and C. C. Kemp, "Active robot-assisted feeding with a general-purpose mobile manipulator: Design, evaluation, and lessons learned," *Robotics and Autonomous Systems*, vol. 124, p. 103344, 2020.

[18] A. Candeias, T. Rhodes, M. Marques, M. Veloso *et al.*, "Vision augmented robot feeding," in *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, 2018.

[19] J. Grannen, Y. Wu, S. Belkhale, and D. Sadigh, "Learning bimanual scooping policies for food acquisition," in *6th Annual Conference on Robot Learning*, 2022.

[20] R. K. Jenamani, D. Stabile, Z. Liu, A. Anwar, K. Dimitropoulou, and T. Bhattacharjee, "Feel the bite: Robot-assisted inside-mouth bite transfer using robust mouth perception and physical interaction-aware control," in *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, 2024, pp. 313–322.

[21] I. Choi, K. Ko, H. Song, B.-J. Jung, J.-H. Hwang, H. Moon, and G.-H. Yang, "A meal-assistance robot system for asian food and its food acquisition point estimation and user interface based on face recognition," *Applied Sciences*, vol. 13, no. 5, p. 3216, 2023.

[22] E. K. Gordon, R. K. Jenamani, A. Nanavati, Z. Liu, D. Stabile, X. Dai, T. Bhattacharjee, T. Schrenk, J. Ko, H. Bolotski *et al.*, "An adaptable, safe, and portable robot-assisted feeding system," in *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, 2024, pp. 74–76.

[23] G. Canal, G. Alenyà, and C. Torras, "Personalization Framework for Adaptive Robotic Feeding Assistance," in *International Conference on Social Robotics (ICSR)*. Springer International Publishing, 11 2016, pp. 22–31.

[24] ——, "Adapting robot task planning to user preferences: an assistive shoe dressing example," *Autonomous Robots*, vol. 43, no. 6, pp. 1343–1356, 2019.

[25] R. Karim, A. Nanavati, T. A. K. Faulkner, and S. S. Srinivasa, "Investigating the levels of autonomy for personalization in assistive robotics," 2023.

[26] J.-A. Yow, N. P. Garg, M. Ramanathan, W. T. Ang *et al.*, "Extract-explainable trajectory corrections for language-based human-robot interaction using textual feature descriptions," *Frontiers in Robotics and AI*, vol. 11, p. 1345693, 2024.

[27] J.-A. Yow and W. T. Ang, "Savr: Scooping adaptation for variable food properties via reinforcement learning," in *2025 International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2025.

[28] S. Liu, Z. Zeng, T. Ren, F. Li, H. Zhang, J. Yang, Q. Jiang, C. Li, J. Yang, H. Su *et al.*, "Grounding dino: Marrying dino with grounded pre-training for open-set object detection," in *European Conference on Computer Vision*. Springer, 2025, pp. 38–55.

[29] T. Khot, H. Trivedi, M. Finlayson, Y. Fu, K. Richardson, P. Clark, and A. Sabharwal, "Decomposed prompting: A modular approach for solving complex tasks," *arXiv preprint arXiv:2210.02406*, 2022.

[30] T. B. Brown, "Language models are few-shot learners," *arXiv preprint arXiv:2005.14165*, 2020.

[31] J. Wei, X. Wang, D. Schuurmans, M. Bosma, F. Xia, E. Chi, Q. V. Le, D. Zhou *et al.*, "Chain-of-thought prompting elicits reasoning in large language models," *Advances in neural information processing systems*, vol. 35, pp. 24 824–24 837, 2022.

[32] H. W. Chang, J.-A. Yow, L. S. Lim, and W. T. Ang, "Design of a breakaway utensil attachment for enhanced safety in robot-assisted feeding," in *2025 International Conference On Rehabilitation Robotics (ICORR)*. IEEE, 2025, pp. 829–834.