Design of a Spoken Dialogue System to Provide Food Knowledge to Users while Eating

Daiki Mayumi
mayumi.daiki.mb9@is.naist.jp
Nara Institute of Science and Technology
Ikoma, Japan

Yugo Nakamura
y-nakamura@ait.kyushu-u.ac.jp
Kyushu University
Fukuoka, Japan

Yuki Matsuda
yukimat@is.naist.jp
Nara Institute of Science and Technology
Ikoma, Japan

Annalena Aicher
annalena.aicher@uni-ulm.de
Ulm University
Ulm, Germany

Keiichi Yasumoto
yasumoto@is.naist.jp
Nara Institute of Science and Technology
Ikoma, Japan

Wolfgang Minker
wolfgang.minker@uni-ulm.de
Ulm University
Ulm, Germany

ABSTRACT

In this paper, we introduce the concept of a spoken dialogue system, which aims to recognize user eating behaviors during meals and promotes healthier eating habits by engaging users in an interactive co-dining experience. To provide an intuitive and more natural access, this information is integrated into a natural conversation with the user. Therefore, we conducted a preliminary study to explore how a human-like virtual avatar's eating speed affects the eating pace of a co-dining user. The results indicate that dining together with an avatar reduces the user's feeling of loneliness and enhances the overall satisfaction and enjoyment of the meal. Therefore, this work represents an important step toward realizing engaging co-dining experiences with human-like avatars and provides new insights into the future design of IoT systems to support solitary dining and health recommendations.

CCS CONCEPTS

• **Human-centered computing** → Natural language interfaces; *User studies.*

KEYWORDS

Human Food Interaction, Eating recognition, Persuasive technology, Intelligent Virtual Avatar

ACM Reference Format:
Daiki Mayumi, Yugo Nakamura, Yuki Matsuda, Annalena Aicher, Keiichi Yasumoto, and Wolfgang Minker. 2023. Design of a Spoken Dialogue System to Provide Food Knowledge to Users while Eating. In *The International Conference on the Internet of Things (IoT 2023), November 7–10, 2023, Nagoya, Japan.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3627050.3630728

IoT 2023, November 7–10, 2023, Nagoya, Japan

ACM ISBN 979-8-4007-0854-1/23/11.
https://doi.org/10.1145/3627050.3630728

1 INTRODUCTION

Sharing meals with someone is an essential aspect of human daily routines. Eating together, referred to as "commensality", imbues individuals with a sense of comfort and thereby stimulates their communicative capabilities. Moreover, commensality plays a pivotal role in strengthening bonds between humans and further supports the development of individual taste preferences [3]. However, due to recent societal influences, such as those brought about by the COVID-19 pandemic, including rising rates of late marriages or decisions to remain unmarried, the proportion of people eating alone ("solitary dining") has noticeably increased. The prevalence of solitary dining has accelerated the diversification of dietary habits, enticing individuals toward unhealthy food choices and eating behaviors. This skewness in dietary habits comes with various disadvantages in terms of health, nutrition, and social interactions [20]. Concerning health drawbacks, there is an increase in potential risks for obesity and metabolic syndrome [10]. Moreover, solitary eating is associated with a higher prevalence of unhealthy eating behaviors [23], and a reduction in meal variety [24]. The domain of Human Food Interaction (HFI) [1, 19] specifically emphasizes reimagining human-food relationships through the augmentation of dining experiences via technological innovations. However, current research in this field focuses on enhancing sensory food experiences through multisensory approaches. However, at current, the majority of research is related to multisensory experience approaches that alter or enhance the sensory experience of food, and there are very few examples of the development of systems that aim to fundamentally change attitudes and establish a healthy diet. In contrast, in our research, we aim to convey healthy dietary knowledge within a natural spoken interaction during a shared meal between an avatar and a human user.

To achieve this, we will construct and evaluate a voice dialogue system that facilitates high-affinity interactions during meals and provides users with valuable information on healthy eating. The proposed system is composed of a mechanism that recognizes eating behaviors and a voice dialogue system furnished with dietary knowledge. By engaging in shared meals through this system, we aim to mitigate the issues associated with solitary dining and enhance users' dietary intake behaviors and their knowledge about

meals. In this paper, we will elucidate the design and implementation of the proposed voice dialogue system and report the preliminary experimental findings regarding the influence that shared meals exert on individuals.

The remainder of the paper is as follows: after an overview of related work in Section 2, Section 3 gives an overview of the proposed dialogue system. In Section 4 and Section 5 our preliminary study and respective results are described, followed by a brief conclusion and outlook on future work in Section 6.

2 RELATED WORK

Research in the domain of Human-Computer Interaction (HCI) related to dining can be categorized into different approaches [26]. The first approach focuses on leveraging various sensory experiences to alter or enhance the taste of food. The second approach focuses on using technology to address the issue of solitary eating among individuals. The third aims to channel individuals towards adopting healthier eating habits.

2.1 Enhancing the Taste Experience

In the field of Human-Computer Interaction (HCI), various methods for enhancing the taste and flavor perception of food have been proposed. These include approaches such as using Galvanic Taste Stimulation (GTS) from both inside and outside the oral cavity to enhance saltiness [15], studies demonstrating the induction of different tastes by electrically stimulating the taste receptor system on the tongue [21], and research altering taste by presenting scents to the olfactory senses [12, 13]. Furthermore, in existing HCI research, the focus has primarily been on visual elements, with methods proposed for manipulating the appearance and visual aspects of food. For instance, systems have been developed to control satiety by altering the quantity of food consumed using Head-Mounted Displays [17]. Similarly, Projection Mapping techniques have been explored to change the visual aspects of food and the dining environment [22].

2.2 Expanding Solitary Dining Experience

Commensality refers to the act of eating together or sharing a meal at the same table [5]. Due to recent societal changes the time people spend dining with family or friends is notably decreasing. Consequently, the phenomenon of solitary dining, has emerged as an issue. Often, eating alone is perceived as not particularly pleasurable [9]. Nonetheless, eating alone has become a familiar routine in modern life, leading to individualized meal patterns and changes of commensality [28]. To address the problem of solitary eating, in the context of HCI, efforts have been made to incorporate communication into the dining experience as a part of digital or computer-enhanced commensality.

For instance, systems have been proposed that bridge the time and spatial distance between diners in remote locations [18, 25, 27]. Numerous interactive dining tables have been introduced, enabling unique interactions with fellow diners [4, 14]. Furthermore, there have been systems designed to create artificial dining companions instead of human ones [7].

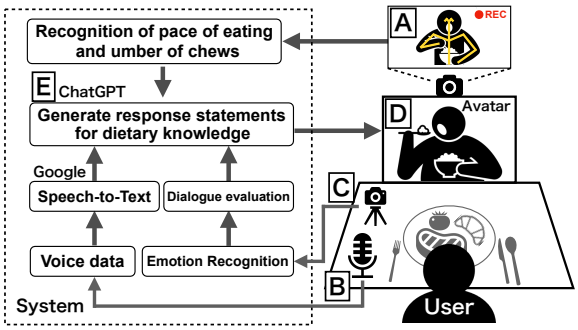

Figure 1: System Overview

2.3 Eating Behavior

For promoting healthier eating habits, behavioral interventions, such as slowing down the speed of eating [2, 11] and increasing the number of chews before swallowing, have been suggested [29]. Nakamura et al. developed the eat2pic system, which leverages visual cues by reframing the act of eating into a task of coloring a picture, thus advocating healthier dietary practices [16]. Moreover, wearable devices like Slowee, which signals users through light and vibrations when they eat too fast, have been proposed [8]. Sensing forks that vibrate to alert users, particularly those who self-identify as fast eaters, aim to raise awareness and regulate eating speed [6]. While there have been studies aimed at enhancing taste and dining experience using information technology, there are very few examples of systems targeting a foundational shift in consciousness and the establishment of a consistently healthy dietary lifestyle.

3 PROPOSED SYSTEM

In order to be able to detect unhealthy eating habits, the system needs to be equipped with the respective sensors and processing capabilities to analyse the user behavior in real-time. Therefore, we aim for a robust design that enhances the accuracy of sensor recognition during food intake, which enables the system to be easy to use in everyday life. Additionally, to provide valid and well-considered health and food information, an integration of an appropriate knowledge source is necessary. In a first step the system shall guide users towards acquiring correct food knowledge, especially to avoid unhealthy food choices and rapid consumption.

In more detail, the proposed system consists of five components: PC, a monitor, a webcam, a depth camera [1] and a microphone. The first, using the depth camera positioned at the top of the monitor, the system captures the user's eating behaviors. This technology allows for the detection of the user's skeletal movements, enabling the system to recognize the meal's pace and chewing frequency without the need for sensors on the tableware (Figure 1, A). The second, voice inputs from the user are captured by a microphone and processed through Google's speech-to-text API (Figure 1, B). The third, facial recognition is performed using a webcam to recognize the user's emotions, the outcomes of which determine whether to maintain the dialogue and if the user acknowledges the content

[1] https://store.intelrealsense.com/buy-intel-realsense-depth-camera-d457.html?-1504367728.1664317294

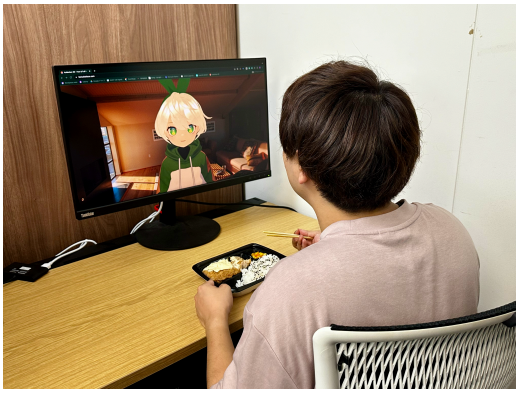

Figure 2: Experiment Scene

of the conversation (Figure 1, C). The fourth, on the screen, an anthropomorphic avatar joins the user in dining. Based on the data obtained from the user's eating speed in part A, the avatar's eating pace is adjusted, effectively guiding the user towards a healthier eating rhythm (Figure 1, D). Finally, the system then crafts a response for the user. By using the recognition outcomes from parts A,B and C, the system, through ChatGPT, generates feedback concerning healthy eating habits (Figure 1, E). This feedback is presented to the user via the avatar shown in Figure 1, D. Furthermore, to optimize the voice recognition rate from Figure 1, B, the system is designed to encourage the user to speak at moments when they aren't chewing, based on the data (Figure 1, A). This approach minimizes potential recognition errors, paving the way for a robust system design. To investigate the impact of the proposed system on dietary behavior and the acquisition of healthy eating knowledge, we conducted pre- and post-experiment surveys on health literacy. Furthermore, we examined the influence of users' nationality and culture and evaluated how information technology interventions may lead to varying outcomes based on user attributes.

4 PRELIMINARY STUDY

To examine whether sharing a meal with an avatar alleviates the feeling of loneliness during solitary dining, we conducted a preliminary experiment. The study involved ten male participants in their twenties (average age = 22.7). A lunch box was provided, and all participants consumed the same meal during each experiment. The experiment consisted of two sessions. In the first session, participants ate alone, engaging in activities such as using their mobile phones or watching videos, simulating their typical dining experience. In the second session, participants shared the meal with an avatar (Figure 2). The avatar engaged in casual conversation, discussing everyday topics, hobbies, and making statements like a human would. Throughout the experiment, the participants' mealtime interactions were recorded with cameras. After the meal the participants were asked to complete a questionnaire. The questionnaire utilized a 7-point Likert scale and consisted of four items (Q1: How enjoyable was your mealtime experience? Q2: How lonely did you feel during the meal? Q3: Were you able to savor the taste of the meal? Q4: How satisfied were you with the meal?).

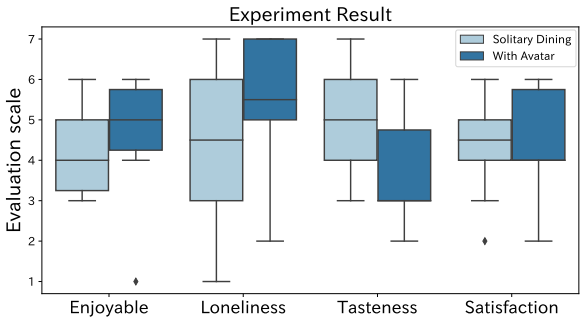

Figure 3: Result of Experiment

5 RESULTS

The results are illustrated in Figure Figure 3. The results indicate that, compared to solitary dining, sharing a meal with an avatar tends to reduce feelings of loneliness and increase satisfaction during the mealtime. However, it was observed that dining alone received higher ratings for the enjoyment of mealtime and the perception of taste. In other words, the practice of dining with someone, even if they are not human, may be an effective method for mitigating feelings of loneliness. Additionally, the decrease in mealtime satisfaction can be attributed to respondents who reported feeling better mentally when eating while conversing with others. Conversely, some users expressed feelings of tension during avatar-mediated dining, stating uncertainty about who was behind the avatar. Others mentioned that they found dining alone to be more mentally enjoyable compared to dining with an avatar.

This experiment has revealed several expectations that users have towards avatars. It was found to be challenging for avatars to evoke the same level of familiarity as humans. On the other hand, users expressed a desire for avatars that can provide new insights into topics related to their interests and hobbies. In this experiment, we employed a format where avatars were questioned about mealtime and hobbies. However, there is a potential need for future experiments to investigate the receptiveness of avatars in providing insights into these topics. Furthermore, since the participants in this study consisted only of males in their twenties, it is essential to broaden the participant pool to include females and individuals of varying age groups in future experiments. Additionally, exploring the impact on individuals who have limited opportunities for daily conversation, such as elderly persons, would be a valuable avenue for further research.

6 CONCLUSION

In this work, we proposed a spoken dialogue system designed to foster high-affinity interactions during meals, particularly during shared dining experiences. One of its main purposes is to address the issue of solitary dining, improve the eating behaviors of users, and provide them with knowledge about healthy eating habits. In an initial experiment we investigated the impact of dining alone versus dining with an avatar. The results indicated that sharing meals reduced feelings of loneliness and increased satisfaction with mealtime. However, the relationship between avatar and human is crucial, as dining with an avatar, who is not perceived as a friend, induced feelings of tension and thus, less pleasure. This valuable

insight into the user expectations concerning co-dining avatars will serve as a basis for designing the system proposed in this work and thus, contribute to the development of future IoT technologies in the context of Human Food Interaction.

ACKNOWLEDGMENT

This study was supported in part by JST PRESTO under Grant No. JPMJPR2039.

REFERENCES

[1] Ferran Altarriba Bertran, Samvid Jhaveri, Rosa Lutz, Katherine Isbister, and Danielle Wilde. 2019. Making Sense of Human-Food Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300908

[2] Dieuwerke P. Bolhuis and Russell S.J. Keast. 2016. Assessment of eating rate and food intake in spoon versus fork users in a laboratory setting. *Food Quality and Preference* 49 (2016), 66–69. https://doi.org/10.1016/j.foodqual.2015.11.015

[3] Carole Counihan. 2018. Commensality and taste. *International Journal of Food Design* 3, 2 (2018), 143–146. https://doi.org/10.1386/ijfd.3.2.143_3

[4] Roelof A. J. de Vries, Gijs H. J. Keizers, Sterre R. van Arum, Juliet A. M. Haarman, Randy Klaassen, Robby W. van Delden, Bert-Jan F. van Beijnum, and Janet H. W. van den Boer. 2021. Multimodal Interactive Dining with the Sensory Interactive Table: Two Use Cases. In *Companion Publication of the 2020 International Conference on Multimodal Interaction* (Virtual Event, Netherlands) *(ICMI '20 Companion)*. Association for Computing Machinery, New York, NY, USA, 332–340. https://doi.org/10.1145/3395035.3425654

[5] Håkan Jönsson, Maxime Michaud, and Nicklas Neuman. 2021. What Is Commensality? A Critical Discussion of an Expanding Research Field. *IJERPH* 18, 12 (June 2021), 1–17.

[6] Azusa Kadomura, Cheng-Yuan Li, Yen-Chang Chen, Koji Tsukada, Itiro Siio, and Hao-hua Chu. 2013. Sensing Fork: Eating Behavior Detection Utensil and Mobile Persuasive Game. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (Paris, France) *(CHI EA '13)*. Association for Computing Machinery, New York, NY, USA, 1551–1556. https://doi.org/10.1145/2468356.2468634

[7] Rohit Ashok Khot, Eshita Sri Arza, Harshitha Kurra, and Yan Wang. 2019. FoBo: Towards Designing a Robotic Companion for Solo Dining. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3290607.3313069

[8] Joohee Kim, Kwang-Jae Lee, Mankyung Lee, Nahyeon Lee, Byung-Chull Bae, Genehee Lee, Juhee Cho, Young Mog Shim, and Jun-Dong Cho. 2016. Slowee: A Smart Eating-Speed Guide System with Light and Vibration Feedback. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 2563–2569. https://doi.org/10.1145/2851581.2892323

[9] Dannie Korsgaard, Thomas Bjøner, and Niels Christian Nilsson. 2019. Where would you like to eat? A formative evaluation of mixed-reality solitary meals in virtual environments for older adults with mobility impairments who live alone. *Food Research International* 117 (2019), 30–39. https://doi.org/10.1016/j.foodres.2018.02.051 Special issue on "Virtual reality and food: Applications in sensory and consumer science".

[10] A Rom Kwon, Yeong Sook Yoon, Kyong Pil Min, Yoon Kyung Lee, and Ji Ho Jeon. 2018. Eating alone and metabolic syndrome: A population-based Korean National Health and Nutrition Examination Survey 2013–2014. *Obesity Research & Clinical Practice* 12, 2 (2018), 146–157. https://doi.org/10.1016/j.orcp.2017.09.002

[11] Sook Ling Leong, Clara Madden, Andrew Gray, Debra Waters, and Caroline Horwath. 2011. Faster Self-Reported Speed of Eating Is Related to Higher Body Mass Index in a Nationwide Survey of Middle-Aged Women. *Journal of the American Dietetic Association* 111, 8 (2011), 1192–1197. https://doi.org/10.1016/j.jada.2011.05.012

[12] Daiki Mayumi, Yugo Nakamura, Yuki Matsuda, Kentaro Ueda, Shinya Misaki, and Keiichi Yasumoto. 2023. Kaolid: Lid-type Olfactory Interface to Improve Taste of Beverages with Ortho-Retronasal Smell. In *2023 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*. IEEE Computer Society, Los Alamitos, CA, USA, 303–305. https://doi.org/10.1109/PerComWorkshops56833.2023.10150412

[13] Daiki Mayumi, Yugo Nakamura, Shinya Misaki, Yuki Matsuda, and Keiichi Yasumoto. 2022. Aromug: Mug-Type Olfactory Interface to Assist in Reducing Sugar Intake. In *Adjunct Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2022 ACM International Symposium on Wearable Computers* (Cambridge, United Kingdom) *(UbiComp/ISWC '22 Adjunct)*. Association for Computing Machinery, 183–187.

[14] Yash Dhanpal Mehta, Rohit Ashok Khot, Rakesh Patibanda, and Florian 'Floyd' Mueller. 2018. Arm-A-Dine: Towards Understanding the Design of Playful Embodied Eating Experiences. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play* (Melbourne, VIC, Australia) *(CHI PLAY '18)*. Association for Computing Machinery, New York, NY, USA, 299–313. https://doi.org/10.1145/3242671.3242710

[15] Hiromi Nakamura, Tomohiro Amemiya, Jun Rekimoto, Hideyuki Ando, and Kazuma Aoyama. 2021. Anodal Galvanic Taste Stimulation to the Chin Enhances Salty Taste of NaCl Water Solution. *Journal of Robotics and Mechatronics* 33, 5 (10 2021), 1128–1134. https://doi.org/10.20965/jrm.2021.p1128

[16] Yugo Nakamura, Rei Nakaoka, Yuki Matsuda, and Keiichi Yasumoto. 2023. Eat2pic: An Eating-Painting Interactive System to Nudge Users into Making Healthier Diet Choices. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 1, Article 24 (mar 2023), 23 pages. https://doi.org/10.1145/3580784

[17] Takuji Narumi, Yuki Ban, Takashi Kajinami, Tomohiro Tanikawa, and Michitaka Hirose. 2012. Augmented Perception of Satiety: Controlling Food Consumption by Changing Apparent Size of Food with Augmented Reality. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 109–118. https://doi.org/10.1145/2207676.2207693

[18] Mamoun Nawahdah and Tomoo Inoue. 2013. Virtually Dining Together in Time-Shifted Environment: KIZUNA Design. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (San Antonio, Texas, USA) *(CSCW '13)*. Association for Computing Machinery, New York, NY, USA, 779–788. https://doi.org/10.1145/2441776.2441863

[19] Anton Nijholt, Carlos Velasco, Marianna Obrist, Katsunori Okajima, and Charles Spence. 2018. 3rd International Workshop on Multisensory Approaches to Human-Food Interaction. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction* (Boulder, CO, USA) *(ICMI '18)*. Association for Computing Machinery, New York, NY, USA, 657–659. https://doi.org/10.1145/3242969.3265860

[20] Barry M Popkin, Linda S Adair, and Shu Wen Ng. 2012. Global nutrition transition and the pandemic of obesity in developing countries. *Nutrition Reviews* 70, 1 (01 2012), 3–21. https://doi.org/10.1111/j.1753-4887.2011.00456.x arXiv:https://academic.oup.com/nutritionreviews/article-pdf/70/1/3/24094379/nutritionreviews70-0003.pdf

[21] Nimesha Ranasinghe, Meetha Nesam James, Michael Gecawicz, Jonathan Bland, and David Smith. 2020. Influence of Electric Taste, Smell, Color, and Thermal Sensory Modalities on the Liking and Mediated Emotions of Virtual Flavor Perception. In *Proceedings of the 2020 International Conference on Multimodal Interaction*. 296–304.

[22] Sho Sakurai, Takuji Narumi, Yuki Ban, Tomohiro Tanikawa, and Michitaka Hirose. 2015. CalibraTable: Tabletop System for Influencing Eating Behavior. In *SIGGRAPH Asia 2015 Emerging Technologies* (Kobe, Japan) *(SA '15)*. Association for Computing Machinery, New York, NY, USA, Article 4, 3 pages. https://doi.org/10.1145/2818466.2818483

[23] Charles Spence. 2017. Comfort food: A review. *International Journal of Gastronomy and Food Science* 9 (2017), 105–109. https://doi.org/10.1016/j.ijgfs.2017.07.001

[24] Wakako Takeda and Melissa K. Melby. 2017. Spatial, temporal, and health associations of eating alone: A cross-cultural analysis of young adults in urban Australia and Japan. *Appetite* 118 (2017), 149–160. https://doi.org/10.1016/j.appet.2017.08.013

[25] Hitomi Tsujita, Svetlana Yarosh, and Gregory D. Abowd. 2010. CU-Later: A Communication System Considering Time Difference. In *Proceedings of the 12th ACM International Conference Adjunct Papers on Ubiquitous Computing - Adjunct* (Copenhagen, Denmark) *(UbiComp '10 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 435–436. https://doi.org/10.1145/1864431.1864474

[26] Carlos Velasco, Qian J Wang, Marianna Obrist, and Anton Nijholt. 2023. A reflection on the state of multisensory human–food interaction research. (6 2023). https://sussex.figshare.com/articles/journal_contribution/A_reflection_on_the_state_of_multisensory_human_food_interaction_research/23495525

[27] Jun Wei, Xuan Wang, Roshan Lalintha Peiris, Yongsoon Choi, Xavier Roman Martinez, Remi Tache, Jeffrey Tzu Kwan Valino Koh, Veronica Halupka, and Adrian David Cheok. 2011. CoDine: An Interactive Multi-Sensory System for Remote Dining. In *Proceedings of the 13th International Conference on Ubiquitous Computing* (Beijing, China) *(UbiComp '11)*. Association for Computing Machinery, New York, NY, USA, 21–30. https://doi.org/10.1145/2030112.2030116

[28] Luke Yates and Alan Warde. 2016. Eating together and eating alone: Meal arrangements in British households. *The British journal of sociology* 68 (11 2016). https://doi.org/10.1111/1468-4446.12231

[29] Yong Zhu and James H. Hollis. 2014. Increasing the Number of Chews before Swallowing Reduces Meal Size in Normal-Weight, Overweight, and Obese Adults. *Journal of the Academy of Nutrition and Dietetics* 114, 6 (2014), 926–931. https://doi.org/10.1016/j.jand.2013.08.020