Food Opera:
A New Genre for Audio-Gustatory Expression

Ben Houge, Berklee College of Music;
Jutta Friedrichs, independent researcher

Abstract

"Food opera" is the term that the authors have applied to a new genre of audio-gustatory experience, in which a multi-course meal is paired with real-time, algorithmically generated music, deployed over a massively multichannel sound system. This paper presents an overview of the system used to deploy the sonic component of these events, while also exploring the history and creative potential of this unique multisensory format.

Introduction

Since 2012, we have produced three audio-gustatory events that we have termed "food operas." At these events, a multicourse meal, designed by Chef Jason Bond of Bondir restaurant in Cambridge, MA, is accompanied by a real-time generative score, composed by Ben Houge. This score responds to events cued by diners, drawing on event-driven scoring techniques adapted from Houge's work in the video game industry, dating back to 1996. Each table in the restaurant is outfitted with speakers, one for each diner, totaling thirty channels of coordinated, real-time, algorithmic, spatially deployed sound in all.

The concept is predicated on the acknowledgement of dining as a time-based art form, akin to film, dance, and music. We use the term "opera" for its multimedia associations, describing a hybrid, interdisciplinary work that engages and explores the junctures between multiple senses.

We identify the main innovations in our work as follows:
- Real-time scoring techniques adapted from the field of video game music create a custom soundtrack that follows unpredictable, live input from diners.
- Associating sound with taste at the level of texture (somewhere between an individual note and a complete musical phrase) allows for a heterophony in which the two senses can be closely linked while still maintaining a degree of autonomy.
- Coordinating sound across a massively multichannel speaker array allows us to preserve the interpersonal and social aspect of communal dining while also elaborating a large-scale sonic environment.

Each of these innovations on its own may represent a small step forward, but collectively, they allow for what we consider to be a new genre of artistic production: the food opera.

Through the food opera project, we hope to show how the application of real-time music generation techniques, originally developed for use in the virtual worlds of video games, can increasingly be applied to score the unpredictable, real-time phenomena of our everyday physical environment, which we observe to be a rapidly growing arena of cultural production.

History and Previous Work

The impetus to associate dining with music goes back to our earliest human history; in myriad historical documents, music has been present at feasts, banquets, rituals, and celebrations.[1] Since the development of the restaurant, performers have often been present to provide background music for diners, and since the advent of recorded sound in the twentieth century, recorded music, from jukeboxes to iPods, has increased in prominence in restaurants to the point of ubiquity. Dinner theaters rose to prominence in the 20th century in the United States. But in all of these situations, music, by the very means of its production, has been physically distanced from the table and at best only coarsely synchronized with the meal, with none of the sophistication of interplay between music and other art forms to which we are accustomed in a ballet, a film, or an opera.

---

[1] For a lively and detailed discussion of historical pairings of food with music, see Qian (Janice) Wang's MIT master's thesis *Music, Mind, and Mouth: Exploring the Interaction Between Music and Flavor Perception* (2013).

In recent years, there has been an increase in artistic attention to the sense of taste, while at the same time chefs have been making forays into the art world. Perhaps the most famous milestone in this rapprochement was Spanish chef Ferran Adrià's *G Pavilion* at the Documenta 12 art exhibition in 2007. Other touchstones include Natalie Jeremijenko and Mihir Desai's *Cross(x)Species Adventure Club* and renowned French pastry chef Pierre Hermé being invited to lecture at the Harvard Graduate School of Design.

Examples of food-based experiences that incorporate sound are more rare. The most well known example is perhaps *The Sound of the Sea*, developed at Heston Blumenthal's Fat Duck restaurant in Bray, England; this dish is served with an iPod allowing diners to hear a field recording of ocean surf while they dine. A similar approach is used in *Volcano Flambé*, developed by chef Kevin Lasko and artist Marina Abramović at Park Avenue Winter in New York, which was accompanied by a recording of the artist describing the ice cream and merengue-based dessert and thanking diners "for eating with awareness." Paul Pairet, a French chef based in Shanghai, opened a new restaurant named Ultraviolet in the summer of 2012, which seats only ten diners per night, and each of the twenty-two courses is accompanied by video projections on the walls and a soundtrack of previously-composed music (the Beatles' "Ob-la-di, Ob-la-da" accompanies a riff on traditional English fish and chips, for example).

While there has been an increase in art world interest in eating in the last century, much of it ignores the sensory nuance of taste. We dismiss as irrelevant to the current discussion art that incorporates food for purely visual or conceptual effect (e.g., Luciano Fabro's *Sisyphus*, Paola Pivi's *It's a Cocktail Party*), performance work that involves eating without an exploration of the sensation of taste (Yoko Ono's *Tunafish Sandwich Piece*, Alison Knowles' *Identical Lunch*, Emily Katrencik's *Consuming 1.956 Inches Each Day For 41 Days*, in which she ate a hole in a Brooklyn gallery wall), or merely visual representations of food (Claes Oldenburg's *Baked Potato*, Francesc Guillamet's photographs of dishes from El Bulli).

Our work has its genesis in conversations and sketches by one of our authors, Ben Houge, dating back to 2006, including a workshop in Shanghai in 2010. Our first event was entitled *Food Opera: Four Asparagus Compositions* and took place in May 2012 at Harvard's Graduate School of Design, as part of a food-oriented series curated by Jutta Friedrichs, Elizabeth MacWillie, and Sara Hendren, students in the GSD's Art, Design and the Public Domain program, headed by professors Sanford Kwinter and Krzysztof Wodiczko. The second event, entitled *Sensing Terroir: A Harvest Food Opera*, took place in November 2012 at Chef Jason Bond's Bondir restaurant in Cam-

bridge, sponsored by Artists in Context and produced by Jutta Friedrichs; this event incorporated field recordings and interviews with local, organic food providers and sought to explore dining as a narrative form to tell the story of sustainable food practices. The third event, entitled *Beside the White Chickens: A Summer Food Opera*, took place in July 2013 at Bondir and highlighted Chef Bond's poultry providers, Pete and Jen's Backyard Birds of Concord, MA, taking as its inspiration the William Carlos Williams poem, "The Red Wheelbarrow."

Video Games in the Restaurant

Any observed event happening in real-time in the physical world is indeterminate, to some extent, from the perspective of the observer. As video games are inherently organized around the unpredictable input of players, video game music has evolved to be uniquely sensitive to real-time input. In a video game, any event may be associated with a musical parameter, and this concept of mapping between two different types of data is at the core of video game music, and indeed, of digital art in general. These mappings have evolved to be quite sophisticated and can be thought of as points of convergence between two autonomous but linked systems.

The challenge of composing music for real-time deployment may be summarized as follows: deciding what to do when an event is received (typically the managing of a musical event or transition), and what to do for the indeterminate amount of time between events (typically some kind of continuous or state-based musical behavior). The simplest solution, and historically the most prevalent, is to simply loop a piece of music indefinitely until an event is registered, at which point the music fades to silence over a certain amount of time, while a new piece of looping music may fade in. The disadvantages to this most simplistic approach, however, are numerous, including tedious repetition and inelegant transitions.

Many more sophisticated techniques have been developed, and while it is beyond the scope of this paper to present a full taxonomy, the techniques that we have incorporated into our food opera work to date include the coordination of musical events along a metric timeline, transposition of MIDI-like note data, and the algorithmic generation of musical melodies, pitch aggregates, rhythms, and phrases. In some aspects, our focus on notes and short phrases recalls the DirectMusic system released by Microsoft in 1999. It also draws from work by aleatoric composers of the mid-twentieth century, including John Cage, Earle Brown, and Christian Wolff. For more information, see Ben Houge's paper *Cell-Based Musical Deployment in Tom Clancy's EndWar*, presented at the First International Conference on Musical Metacreation in 2012.

In our food opera work to date, the sources of real time events are constrained to a fairly limited range: we observe which dishes are chosen by a diner, when each course arrives, and when it is finished. The number of courses is fixed (four for our first food opera, five for the second and third), and for most courses, the diner has a choice between two dishes.

Each course is associated with a musical texture. The musical texture is not a linear piece of music that plays from beginning to end. Rather, since each course lasts an indeterminate amount of time, depending on a number of dining parameters (most prominently, each diner's rate of food consumption), the texture consists of one or more algorithmic musical behaviors, capable of continuing indefinitely, until an event is received indicating that the behavior should end.

In addition to the events indicating the beginning and ending of each course, the music responds to the elapsed time since the beginning of each course. In most cases, over the course of each course, there is a gradual decrease in musical intensity, which may be quantified in density of musical events, volume, or number of concurrent layers. An additional behavior, limited in duration, may be associated with the beginning or ending of each course.

These parameters may seem few, but given that diners are coming and going throughout a span of roughly five hours, and that any of the twenty-six diners may be at any point in their meal at any time, the richness and complexity of the overall, emergent sonic environment is significant.

Texture and Musical Granularity

What we mean by musical texture is a convenient middle ground between individual note and fully composed musical composition. It is a musical behavior with a clear identity, characterized by instrumentation, melodic contour, phrase duration, harmonic structure, voicing style, rhythmic density, and number of constituent layers. While textures may contain melodic elements, they are not primarily melodic in character. Our textures are designed to continue indefinitely, using algorithmic techniques to vary, shuffle, offset, juxtapose, and in some cases generate musical material.

All game music may be positioned on a continuum of granularity. Sample-based music may be categorized according to the duration of its constituent wave files, from through-composed pieces of several minutes in duration to individual notes (or even, in considering real-time synthesis systems, subdivisions of notes). Granularity is linked to responsiveness, which is to say, the maximum amount of

time that could elapse between the arrival of a game event and a musical system's response. By orienting music around texture, we achieve a highly responsive musical fabric.

We emphasize texture, because it offers a useful way of defining a musical identity without large-scale temporal expectation; a key aspect of our use of texture is that it can continue indefinitely. It is an additive process, allowing the steady accretion of musical material over time. This allows for sustained and concentrated evaluation, which in turn makes it well suited to pairing with other sensory input, such as taste.

We find this approach preferable to playing a completed piece of music, which may contain a conflicting or distracting dramatic trajectory (manifested, for example, as the climax of a phrase or a juncture in a piece's formal articulation), or which may end prematurely, or which may loop unvaryingly, resulting in fatigue or annoyance for listeners due to unmitigated repetition.

This approach is also preferable to simply sustaining or repeating a single tone, timbre, or sonority, as it provides rhythmic and harmonic context for individual tones and allows for the investigation of the modes of meaning with which music has long been associated, including harmonies, rhythms, and melodies, the syntax and grammar of music.

Coordination and Modularity

The music for the food opera is coordinated in harmony and rhythm and composed so that this coordination would be readily apparent. As most of our diners are assumed to be musical non-specialists, we decided to compose music that is diatonic and organized around a clear beat referential.

All rhythmic activity is coordinated in reference to a common pulse of 188 beats per minute that is present throughout the work. New events are, for the most part, quantized to start on a beat (although random delays of a few milliseconds may be introduced as a humanizing gesture). There is no concept of meter, only pulse, although events may be assigned to happen on multiples of beats, which creates a kind of local meter.

As the project has evolved over our three events so far, we determined a desire to blur the rhythm at certain points, to avoid an overly metronomic effect. For this reason we introduced the idea of rubato phrases that could be played intermittently, uncoordinated with the underlying pulse.

Over the course of the evening, the underlying harmony modulates very slowly in a random walk among five diatonic key areas excerpted from the circle of fifths: D, G, C, F, and B flat. This provides a sense of harmonic progression and combats key fatigue. A bass drone (generated in real-time from transposed recordings of a viola, a human voice, a synthesizer, and an electric generator used in a chicken slaughtering we observed) articulates this harmonic movement. The bass drone slowly (every forty to eighty seconds) choses scale degrees from the current tonality (excluding the leading tone) according to a first order Markov chain of acceptable progressions; this allows each texture to be heard with any of six possible scale degrees of the current key in the bass, providing additional variation. Key modulations occur every three to eight bass notes.

Our musical textures are transposed according to one of two principles. For musical textures that are rendered using a MIDI-like system, transposition is a trivial affair, involving an integer semitone offset. For textures that incorporate recorded phrases, transposition is slightly more complicated: we choose from a pre-compiled list representing the subset of phrases compatible with the current key. In some cases, we perform real-time transposition of musical phrases. For rubato phrases that do not need to be rhythmically precise, we have found it musically acceptable to transpose as much as two semitones up or down to achieve compatibility with the current key. For phrases that do require rhythmic precision, we may transpose in simple multiples of the original speed (e.g., 0.5, 0.667, 1.5, 2.0) to achieve musical variety and to maximize the contexts in which a recorded phrase may be used (an important concept in game audio development).

As for the musical textures themselves, each can be thought of as an independent algorithmic study unto itself. The only constraint is that it must in some way conform to the current key area and acknowledge the underlying pulse. Textures are composed to be somewhat sparse, with an awareness that they will be deployed alongside other textures and are designed with registral variation in mind.

The system we used for the first three food operas was programmed in Max. The menu was preloaded into the system, and as diners placed their orders, we input the choices into a patch representing the layout of the restaurant. As dishes came and went, we advanced the musical behavior for each seat. An interstitial behavior, consisting of footsteps from one of the farms providing the ingredients for the meal, quantized to the pulse of the music, played between courses. Ambient sound from the farms played before the first course and after the last.

Renderings of excerpts from the food operas may be heard here:

https://soundcloud.com/benhouge/sets/food-opera-excerpts

Space and Massively Multichannel Sound

An overriding goal throughout this project has been to respect the integrity and the history of the experience of communal dining: the focus is on the table, not on a stage situated elsewhere in the dining environment. Diners are free to converse during the event. To quote Brian Eno, the music should be "as interesting as it is ignorable," a statement that also parallels Chef Jason Bond's observation of the manner in which people enjoy his food. The sound is there to enhance and supplement the meal, not to supersede or distract from it.

In our conception of food opera, the very physical presence of a live performer alters the calculus of the dining experience to an unacceptable degree; this is why we rely on recorded sound for our event. However, because the sound is being deployed in real time, it exists, as all video game soundtracks do, somewhere between recording and performance. The combination of video game music techniques and multichannel sound diffusion are the technologies that enable food opera as we define it to exist.

In our second and third events at Bondir restaurant, each diner has a dedicated speaker through which she or he hears the accompaniment to her or his meal. The small speaker, about 3 inches in diameter, takes the place of a traditional table centerpiece (e.g., a vase of flowers) and rests in a small speaker stand designed by Jutta Friedrichs, which allows two speakers pointed in opposite directions to be directed at two diners facing each other. Bondir seats twenty-six diners, so twenty-six channels of sound are deployed simultaneously to individual diners throughout the restaurant. Six additional speakers (four around the perimeter of the room plus two custom designed six-channel arrays, designed by Stephan Moore, positioned on stands in the middle of the room) play the slowly evolving low frequency drone that provides a harmonic context for the musical textures of each diner's discrete accompaniment.

We do not use any particularly focused speaker technology, and in fact, it is not particularly desirable for our planned experience. Sounds from nearby speakers will naturally blend together, providing a rich field of music, in effect transforming the entire restaurant into a generative sound installation. At our second and third events, diners could make reservations anytime between 5pm and 9:30pm, such that, at any given time, different people in the restaurant would be at different points in their meals; this is why the harmonic and rhythmic coordination described above is particularly important. The overall form of the piece is an aggregate of every table's individual dining trajectory. When diners at different seats in the restau-

rant order the same dish at different times, a kind of emergent foreshadowing and recapitulation occurs that is a unique formal characteristic of this work.

The primary sound sources in our first three events are traditional, acoustic musical instruments. This may seem an unnecessary restriction, as the sound is played through speakers, allowing any electroacoustic sound to be employed. We use traditional musical instrument sounds to create an association with the alchemy of cooking: recognizable physical materials are juxtaposed, manipulated, and transformed to an aesthetic end. There is an essential abstractness that is shared by food and music (distinct from visual art and writing, for example), which we sought to underscore. It is difficult for the taste of a cucumber, like the sound of a violin, to be "about" anything other than itself. As in dining, mimesis in acoustic music is by far the exception (e.g., the timpani thunder rolls at the end of Berlioz's "Scène aux Champs" from *Symphonie Fantastique*). This is intentionally distinct from the associative, mimetic sound world of an electro-acoustic piece such as Paul Lansky's *Table's Clear* (1992), for example, which is based on recordings of kitchen implements.

Future Work

One reason we consider food opera to be a new genre is that its expressive potentialities are so vast.

From an aesthetic perspective, the food opera lends itself to a range of themes, from theatrical narrative structures (including non-linear or generative narratives) that might tell a story, to spatial or landscape meditations that might resemble a sound installation, to ritual situations such as a Passover Seder or a wedding ceremony. The potential to link sound to food, scent, and the tactile sensations of the mouth creates an entirely new field of sensory interplay, which may be harnessed to a wide range of expressive ends.

As an example, in our second and third food operas, alongside the purely musical elements described above, we incorporated field recordings and interviews from some of the local, organic farms that provide Bondir's ingredients. Our goal was to tell the story of sustainable food practices by sonically connecting diners to the sources of their meals. We feel that this message was conveyed more deeply and memorably than it would have been using sound or taste alone.

On the technical side, there is enormous potential to increase the amount of input into the system generating the music; we plan to explore this potential in future food operas via a range of sensing mechanisms. Accelerometers could be attached to stemware or silverware, or cameras could be mounted above the tables, to name only the most obvious subsequent steps. In some situations, it might be desirable to have live musicians performing remotely, similar to what is currently done in some Broadway productions, incorporating real-time scoring and cueing systems.

There is also considerable potential to expand the concept in terms of visual art, incorporating projections and video displays, by coordinating table settings, clothing, and other elements of interior and set design.

References

Wang, Qian (Janice). 2013. Music, Mind, and Mouth: Exploring the Interaction Between Music and Flavor Perception. M.S. diss., Department of Media Arts and Sciences, School of Architecture and Planning, Massachusetts Institute of Technology, Cambridge, MA.