OVERBAKED & UNDERPROOFED - FINAL REPORT

DHPRAXIS SPRING 2023

Team

Maria: project manager & text analyst

Nuraly: researcher & Bingo game designer

RC: web developer & graphic designer

Teddy: data manager & linguist

Project Narrative & Environmental Scan

Overbaked & Underproofed (Ob&Up) offers a closer look at the language used in the judging

segments of *The Great British Baking Show (GBBS)*. *GBBS* is a reality TV baking competition

that has been on-air for more than a decade and has been one of [the most-streamed original tv-](#)

[shows in the US during the pandemic.](#) As the only reality TV show in the top 15, it placed third

in the "original content" category, only surpassed by scripted shows, *Lucifer* and *Squid Game,* in

2021. *GBBS* accumulated over 13 billion viewing minutes on streaming platforms in 2021.

GBBS, *The Great British Baking Show*, as it is known in the US (in Britain, it is known as *The*

Great British Bake-Off), has produced 13 seasons and various holiday-themed specials since its

inception in 2010. It distinguishes itself from other reality shows in several ways: E.g., it is about

finding GB's most skilled amateur baker, and the show's focus is on the "bakes" the contestants

produce and not on their interpersonal intrigues. There is no (deliberate) scheming, and there are no alliances between contestants that result in calculated expulsions. The show's most defining and comforting feature is, in fact, a subversion of the going reality competition mantra "I'm not here to make friends." Contestants of *GBBS* absolutely seem to be there to make friends and help each other through the pressure of the baking challenges.

Nonetheless, it is a competition. One baker must win, and others must be ranked. Judging, under these circumstances, plays an unusual role. Judging is essential to keeping the suspense created by competition alive. The necessary judgment is supplied by two bonafide British experts, the blunt and no-nonsense Paul Hollywood (a British celebrity chef) and Prue Leith (a multi-talented restaurateur, chef, and writer). Every episode gives the judges three opportunities to evaluate bakes. They judge the baker's "signature bake," the "technical challenge," and the "showstopper." In the show, the "signature bake" and the "showstopper" challenges are often introduced via brief segments that feature drawings/renderings of the planned bake along with short stories of the bakers' motivations behind choosing to create this particular bake. The scripted text that supports these visual renderings is narrated by one of the hosts in a lilting, appetite-inducing voice. These almost tender previews suggest that the show's creators know that, in a screen-mediated context, the totality of a bake (taste and all) is transmitted by language as well as visual clues.

 In contrast to the preview, the judging portions seem gruff and linguistically anemic. The judges swiftly move from bake to bake, taste a bite, and immediately pronounce a verdict along a limited list of expressions, like "very good,"' tastes lovely," "absolutely delicious," "good flavor." These fundamentals are at times complemented by a few choice descriptors: "Over-baked," "under-proofed," or vice-versa, "soggy bottom," and "stodgy," and when a contestant's bake

reaches the superlative, language seems to drop out entirely and is replaced by a silent Paul Hollywood Handshake.

Ob&Up set out to explore how language works in the show's judging segments, what language analysis reveals about the use of evaluative language in our culture, and how we relate linguistically to multi-sensory experiences (like food). We worked with an anchor corpus based on Season 12 (which aired during the pandemic, the season described in the opening paragraph). The judging segments of Season 12 were divided into several subcorpora and served as the basis for testing assertions about the size and variety of judging language and vocabulary. Additionally, we created a *GBBS* comparison corpus of Season 2. The comparison corpus let us look at how vocabulary and language use have changed over the course of a decade.

Stylistically and philosophically, *Ob&Up* believes in mixing an academic approach with a playful one. The findings on the project's website are accompanied by a watch-party bingo game that viewers of the show can play while re/watching *GBBS* episodes and a judgment generator — both built on and with judging language. By playfully guiding viewers to notice how judging language contributes to (or reduces) our understanding of food, *Ob&Up* aims to open and continue a conversation about language and sensory experience in our virtual living spaces.

*

When searching for DH projects that might align with *Ob&Up* in television and food culture, it becomes evident that language around food is not often paired with text analysis. Most projects interested in language and food focus primarily on creating archives. The focus is on collecting and archiving recipes and perhaps tying them to specific regions and periods.*[What America](#)*

Ate is an example that uses an archive of 1930 recipes and contextual information to create a multifaceted look at a historical period. Text analysis is not part of the project's goal.

An endeavor that does dive into linguistic aspects of food works with an unusual corpus extracted from menus is *Linguistic Markers of Status in Food Culture: Bourdieu's Distinction in a Menu Corpus*. It compares and analyzes language use on menus of restaurants in various price classes. However, the results are presented as a traditional academic paper. It does include visualizations as infographics but generally follows academic literary conventions. In this way, *Linguistic Markers of Status in Food Culture* seems representative of the prevalent methods for DH projects that employ distant reading in linguistic contexts intending to produce literary, cultural criticism. The focus seems to be EITHER on creating interactive visualization OR on interpreting aspects through distant reading, whereby the visualization created by distant reading tools is used to illustrate findings.

Media like television and radio are not often used to extract corpora and are more likely to serve as an impetus to create archives and collect metadata. One exception is "Oh My God: The linguistic influence of the TV Series Friends" by Garcia, Metzler, Lasser, Pellert, & Di Natale. These authors' work looks at the influence of the TV show *Friends* on the meanings and increased usage of the phrase "Oh My God" in scripted and unscripted television. The project's focus parallels *Ob&Up's* interest in the context and influence of less scripted expressions in culturally dominant TV shows. The project's presentation largely follows academic convention, but it structures its findings in smaller segments that are headed by Friends-style episode titles and inspired *Ob&Up* in that way.

Engagement with the language of television happens more often in a smaller scope and more public forum of media criticism. The Wall Street Journal's recent ["How Much Does John Wick Actually Say? We Counted the Words"](#) by John Jurgensen is an example of the presence of text analysis in public discourse. Much like Jurgensen's article, *Ob&Up's* approach strives to directly reach the public that consumes the media we scrutinize and analyze.

Ob&Up connects and combines the methods of text analysis and metadata collection with close reading to foster media literacy and produce cultural criticism. A. Flicker et al. advocated for this combination in 2018 in the foreword of an editorial special-issue on audiovisual data in DH of *VIEW, Journal of European Television History & Culture*:

"… 21st-century researchers should be encouraged to develop new skills in both close and distant reading techniques: new artful practices of "scalable reading," critical combinations of "explorative" distant listening and viewing, conjoined with "interpretative modes" of close inspection, and so forth. These adaptive skills to zoom in and out between big data and distinctive expressive nuance will serve as an unquestionably challenging yet copiously generative mandate for many years of rigorous research to come." (Flicker et al.)

The quote also points to a larger discourse around DH and cultural criticism that *Ob&Up* wants to enter in a new way. How does DH use its "copiously generative mandate" to produce and further cultural criticism? Alan Liu's essay "Where is Cultural Criticism in the Digital Humanities" speaks to this more significant point. Liu criticizes the disciple as the "practicing partner of distant reading" (Liu) and points to its propensity to overvalue instrumentality and undervalue direct engagement with the public and its experience of culture. For DH to find a constructive place, it needs to work to engage the public in a discourse on culture and the public

life it experiences. In a quote that makes use of reality TV, he frames his argument in this way: "…millions tune in each week to watch crab fishermen on the Discovery Channel (*Deadliest Catch*). Humanists may not be salt-of-the-earth crabbers, and archives may not be as stormy as the high seas. But surely, humanists ought on occasion try to share the excitement of the chase by which breakthrough intellectual discoveries and movements occur" (Liu). With Liu delivering his point by referencing a popular reality TV show, he illustrates the essence that *Ob&Up* seeks to embody.

Audience

The target audience for this project is two-fold. Any (more than casual) fans of *GBBS* —*e.g.*, people who have [shared memes](#) or caught themselves pronouncing a random piece of pastry "stodgy" in a perhaps fake British accent— should find *Ob&Up* intriguing, as should linguistics and other linguistically interested academics in media studies, literature, and beyond.

These two groups are, of course, not mutually exclusive and are likely already involved in ongoing exchanges. Existing language-focused public engagement via word games and puzzles, language-usage quizzes, and the like shows that the channels of communication between the media-consuming public and language experts are naturally present. *Ob&Up* wants to consciously build on overlaps between playful engagement with language and academic research by reaching and connecting these two differently motivated curiosities with an analysis of GBBS judging language communicated, partially, in game form.

Project Activities

Overall, the project integrated several components (corpora and literary analysis, website development, and game components) that each developed in phases. Because the components depended on each other, we had to create a work plan that would consider what phases could develop concurrently and independently and when milestones in one area were crucial to progress in another. Our initial work plan accounted for most of these interdependencies and milestones. To meet milestones and honor the dependencies that make the overall project possible, we had to adjust the scope and expectations for each component in the process.

First and central to the project was, of course, collecting and preparing our data, our corpora.

CORPORA

Initially, we had more ambitious plans regarding how many different shows we were capable of comparing to each other. However, we quickly had to change our priorities and hone in on more specific elements of two GBBS seasons when the captions' transcripts turned out to be less usable than we had hoped. We had planned to compare the seasons of GBBS to other American baking or cooking shows, but it was challenging to find captions for shows that encompassed whole seasons, and even if we had found preexisting scripts in semi-decent shape, readying the corpora for analysis would have required several additional weeks.

The more results we got from analyzing the corpus, the more we had to adapt our direction for further analysis, leading to the inclusion of additional text analysis programs and methods. Initially, our project was more interested in the odd phrases, the memetic phrases that are used within the show, until it turned out that those are the rarer phrases and hence why they become

memetic; the bulk of the language was composed of such unmemorable pronouncements that those phrases stood out.

The preparation of the data, the cleaning of the corpora, was a true team effort. After finding the script of subtitles, RC accomplished a first cleaning via a Text Python Cleaning Script: A Python script (jupyter notebook) to remove "ambient sound" in the transcript was created.

1) Removing Caps

```
# identify all text in all caps
all_caps = set(re.findall(r"(\b(?:[A-Z]+[A-Z]*|[A-Z]*[A-Z]+)\b(?:\s+(?:[A-Z]+[a-z]?[A-Z]*|[A-Z]*[a-z]?[A-Z]+)\b)*)"
                          , textfile
                          , re.DOTALL))

all_caps
```

```
{'A',
 'ALL',
 'BLEEP',
 'BLOWS AIR THROUGH LIPS',
 'CHIGS LAUGHS',
 'DIY',
 'DOG WHIMPERS',
 'DUCK QUACKS\n\nBIRDSONG AND DUCK QUACKS\n\nDRUMSTICKS TAP\n\nMUSIC',
 'GASPS\n\nOh',
 'GROANING',
 'HR',
 'I',
 'IT',
 'LAUGHTER',
 'LAUGHTER\n\nAw',
 'LAUGHTER\n\nI',
 'LAUGHTER It',
 'LINE RINGS',
 'MUSIC CONTINUES\n\nBIRDSONG',
 'NOEL',
 'OK',
 'PAUL',
 'PAUL CHUCKLES\n\nI',
 'PRUE',
 'PRUE MUMBLES'.
```

We then double-checked the pre-cleaned script against the show's spoken language. In the process, we separated the judging sections. For our purposes, the *judging sections* are defined as direct conversations between the judges and contestants in the show's evaluation segments. Within the judging sections, we extracted the judges' language (words spoken by the judges) and then further coded all spoken lines by individual speaker/judge (i.e., Paul or Prue). This process created various sub-corpora for Season 12: Full Season, Judging Segments, Judges Language in Judging Segments, Paul's Judging Language, and Prue's Judging Language.

Additionally, we created a small corpus of baking challenge instruction language (a.k.a. assignment prompts) in Season 12, which we used to facilitate a comparison between assignment language and judging language. For possible comparative queries, we also prepared a corpus for Season 2, the earliest available season that substantially adheres to the show's defining characteristics after a few less enduring ideas had been eliminated after Season 1. Season 2's transcript was also available online, and we followed the same cleaning procedure, although we did not separate the judging language by speaker. And one caveat: Likely for technical reasons, Episode 4 of Season 2 was not available for viewing and had to be accepted as is. The xls files that house the corpora and various subcorpora are available via our GitHub repository.

TEXT ANALYSIS

Once we had created our corpora, we predominately worked with two available online tools to analyze our data.

Voyant is a free and openly accessible online tool showing statistical and visual text analysis results. Because Voyant offers a great variety of visualizations that make it possible to grasp findings easily, we relied on it to discover larger patterns. We ran the corpora with various stopword lists (a list that omits specified words from analysis). We sometimes used the default stopword list and, at other times, manually adjusted the list. Occasionally we test-ran a corpus without a stopword list. All word cloud illustrations shown in our findings have been derived from Voyant. Jason Davies' Wordtree was especially helpful in depicting the complete corpus (without stopwords) in relation to the occurrence of one word or one specific phrase. Illustrations that show branched text were created with Wordtree. Additionally, we also consulted #LancsBox,

language analysis software developed by the University of Lancaster, to add findings to the Season 12—Season 2 comparisons.

We approached text analysis with an open mind and without a particular strategy. We first needed to understand whether what we had observed in single instances, i.e., by closely watching the show, did, in fact, occur. Was there an observable reduction in language? We then wanted to know what words were responsible for the initially unexciting wordscape we had perceived. While looking into the corpus without stopwords, we also learned that the sentence structure contributes to the simplicity we had noted. Experiments in Wordtree showed how benign and unexciting words (like "nice" and "lovely") are at the core of a majority of judgment sentences and how other, more context-specific words galvanize the linguistic identity of the show.

In the process of collecting analysis findings, additional ideas, e.g., for the creation of a judgment barometer, began to emerge. We adjusted the concept of the website to accommodate new ideas. We knew early on that text analysis would be crucial for determining the site's content and that we would have to stay flexible in the later stages of the project. Our findings emerged in increments, and the larger project narrative emerged along with it. To reflect our experience of incremental discovery for an audience, we resolved to share the findings in a blog format, which turned what might have been a drawback into a feature.

WEBSITE

The nature of our process also dictated the development of the site. We knew the development would demand a two-phase approach. We were only able to roughly conceptualize the site prior to text analysis. We employed graphic design best practices and generated a logo, a brand book,

and the website's basic structure. We discussed ways to integrate games and, in tandem with the development of the Bingo game, we decided on a hosting platform that would best facilitate a game tie-in.

Once the text analysis findings became more apparent, we made final decisions on the website structure and experimented with the landing page and the plug-ins. We also added new graphic design elements to the blog posts that feature text analysis findings—the final version of the site aligned with (and exceeded) our initial goals. We are now also in a position to clearly see potential avenues for concrete next steps to expand Ob&Up.

INTERACTIVE GAMES

Our initial plan was to feature the Bingo game as the central focus of the landing page. Here too, we knew that coding the game and considering its implementation on the site would be feasible in the first phase of the project and that we would have to finalize the game post-text analysis and leave extra time for testing the game.

The initial ambitions were clear, but some of the goals became out of reach due to our limited knowledge of Python and Javascript. We were able to make the game work by using a four-year-old Plug-in made for WordPress by another developer, who had updated it last in 2019. Since then, it had been dormant, but it was the only version we could consider since other existing plug-ins were too complicated to implement.

After the initial version of the game was completed, added to the site, and working in a rudimentary way, we could not make improvements and proceed to the testing phase. As a result, we supplemented the site's gaming component with the addition of a judgment generator. This

simple name generator reshuffles segments of judging language in new and often absurd ways. The adjustment helped us maintain the planned focus on games and added a playful way of analyzing language to the project.

Accomplishments

Overbaked & Underproofed is now a website featuring two interactive game as well as our text analysis findings. More specifically, we have developed a WordPress site hosting information about the background (data collection, corpus creation, analysis methods, key findings, a Bingo game, and a judgment generator). We have reached our goal of making a minimum viable product for our website (with some room for improvement surrounding interactive games' embedding).

The entire site was conceptualized to embody part of its mission. *Ob&Up* questions the way we relate to food in a screen-mediated environment and especially wants to point to our privileging of the visual aspects of food at the expense of other senses. To undermine this reliance on the visual, we wanted to strip any images of cakes from the design and create a site that features predominantly verbal language. The landing page shows the result. We center the game components to counteract this refusal of images and signal that our project is for an audience of *GBBS* fans as well as linguistically interested non-*GBBS* fans. Since the games were built on and made with *GBBS* judging phrases and words, they embody our playful approach to the subject matter.

OVERBAKED **&**
UNDERPROOFED

An Exploration of the Great British Baking Show's Judging Language

Overbaked & Underproofed is curious about how language works to widen or shrink an experience of food.

What kinds of words do the judges use to evaluate the contestants' bakes?

How frequently do signature words like "overbaked" & "underproofed" & "stodgy" & "soggy bottom" show up? How does this judging language work?

If you are the kind of person who could be wondering about these things, we have answers! We use text analysis (and a Bingo game) to explore the mechanics and intricacies of GBBS's judging language.

| B	I	N	G	O
Neat as a pin	Under-proofed (More proofing, less baking)	I don't love it	Flavour's coming through beautifully	Absolutely delicious
Absolutely lovely	soggy bottom	Boozy!	The texture's spot-on	Crisp on the outside
Very, very good (double very)	They do look very nice	FREE	Absolutely beautiful	Over-baked
Needed a bit longer in the oven (underbaked)	well baked	An absolute triumph	Shame it looks the way it does	Quite stodgy
It looks a mess!	A bit of a disaster	I'm not getting this [insert flavour]	Good lamination	It looks fantastic

Enter Bingo Game

The Bingo game on the site is easily accessible and playable, as is the judgment generator. The Bingo game opens on a new page and can be played online while watching an episode of the *GBBS*. Since it is populated with extremely frequent to less frequent judging phrases, it invites viewers to listen closely to the judging language on the show.

B	I	N	G	O
Neat as a pin	under-proofed (More proofing, less baking)	I don't love it	Flavour's coming through beautifully	Absolutely delicious
Absolutely lovely	soggy bottom	Boozy!	The texture's spot-on	Crisp on the outside
Very, very good (double very)	They do look very nice	FREE	Absolutely beautiful	Over-baked
Needed a bit longer in the oven (underbaked)	well baked	An absolute triumph	Shame it looks the way it does	Quite stodgy
It looks a mess!	A bit of a disaster	I'm not getting this [insert flavour]	Good lamination	It looks fantastic

Enter Bingo Game

The judgment generator, which also opens to its own window, re-invents judging language by following the basic grammatical structure of a judgment phrase and re-organizing the parts to create subtle breakages and absurdisms. Users can remix the options by pressing the *go back and judge it again* link.
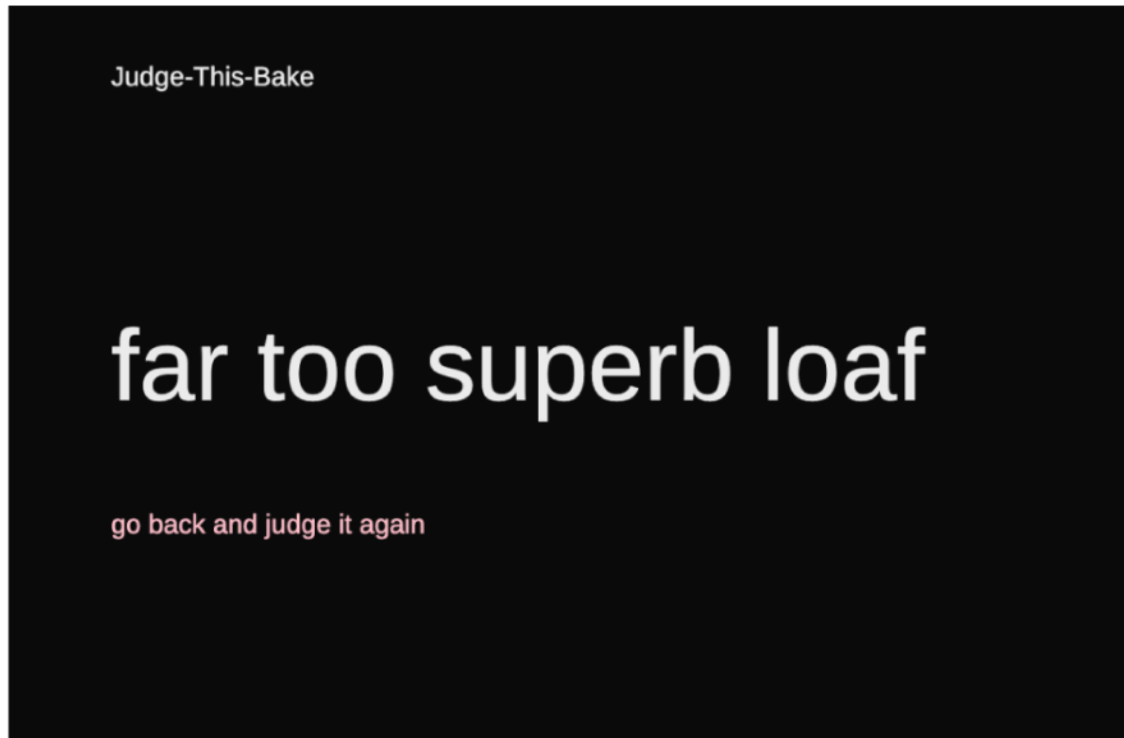
In addition to the games, ten blog posts on the findings page share our text analysis findings. The findings can be accessed sequentially as well as by interest. In addition to text analysis explanations and theories about our findings' meaning and impact, all blog posts feature graphics created by text analysis software programs or, like the judging barometer, by us.
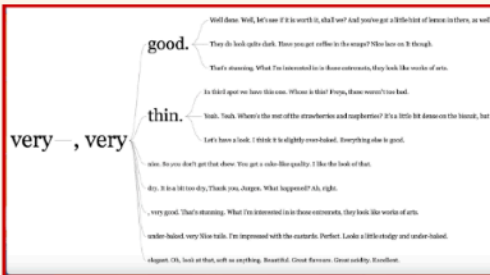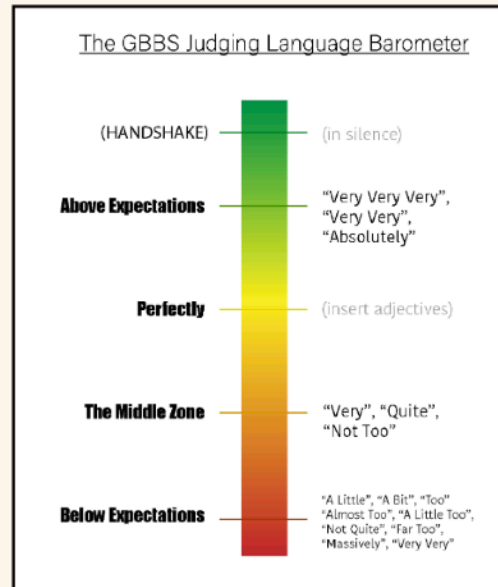
The Sentence Structure of a Judgment

The Bland Canvas

| PERFECT [30] |
| GREAT [38]* |
| BEAUTIFUL [59] |
| DELICIOUS [70] |
| LOVELY [92] |
| GOOD [146] |
| NICE [125] |

Creating Scale

How Is This Show So Nice?

The GBBS Judging Language Barometer

(HANDSHAKE)	(in silence)
Above Expectations	"Very Very Very", "Very Very", "Absolutely"
Perfectly	(insert adjectives)
The Middle Zone	"Very", "Quite", "Not Too"
Below Expectations	"A Little", "A Bit", "Too" "Almost Too", "A Little Too", "Not Quite", "Far Too", "Massively", "Very Very"

Finally, the website also offers insight into the project's process. We include an account of our analysis methods, trace the creation of our corpora, and include a link to a GitHub repository of our data.

16

Evaluation

The project invited and received invaluable feedback from potential audience members, especially in certain aspects. First and foremost, as we created corpora for our team (for each other), we consistently had to think about ways to make the texts usable for queries. Internal exchanges and ideas for refinement, as well as input received through casual conversations with potential audience members and classmates, influenced our way of creating subcorpora. E.g., Wanting to address repeated questions about our claim of the paucity of judging language and questions about what we mean by paucity led to the creation of a small subcorpus of baking assignment language. Using this small corpus to illustrate the reductions of language between assignment and judgment language helped clarify the framework of the project as well as our research questions.

As we developed the findings, we worked closely with several colleagues of Maria's at the Columbia University Writing Center. Also, we consulted the CUNY GC Writing Center and asked for comments. In conversation, we recognized which claims and conclusions of our analysis needed additional evidence or needed to be more explicitly connected. We were able to address questions and integrate feedback to develop the final narrative of our analysis. Through these conversations, we also understood that our findings could be presented in a looser form and that we did not have to present our analysis as a traditional paper. Also aided by WordPress' affordances, the blog format we settled on helped present our analysis to a wider/less exclusive readership in smaller segments that do not have to be read sequentially.

The development of the website also was influenced by feedback from colleagues and internal discussions. As we created the logo and color scheme for the project, especially Teddy's notes on accessibility shaped the paired down and higher-contrast final version. We had decided early on that the website would not feature photographs of cakes and other expected images associated with baking, and so we had to be conscious about not overwhelming a user with text. Centering the games on the landing page mitigated this concern.

When choosing a platform that could accommodate our project, we had to consider our game plug-ins. We needed a platform that would allow for integrating our Bingo game. Finally, a paid and professional WordPress site seemed the most promising. Although it theoretically should house a plug-in, we might have achieved the same result with the CUNY Commons version of WordPress. In order to take full advantage of the paid platform, we might have reached out for additional support in testing and adjusting the game plug-ins. It seemed deceptively simple, and perhaps we were too naive (our team skills were a bit lopsided in this area) to foresee potential frictions. Finally, the professional WordPress site's cost also impacts the project's sustainability.

We tested the game components' functionality with a small circle of potential users. We had scheduled a more extensive testing phase with a relevant *GBBS* Reddit community for the Bingo game. However, we had to delay this effort, partially because of design snags and partially because offering it for testing when a new season of the show airs will be more productive. Until then, we hope to find a way to address the design flaws. Currently, the Bingo game plug-in cannot reset/regenerate (meaning every player plays with the same card every time), and the look & feel of the Bingo page also undercuts the aesthetics of the rest of the website, resulting in a

less than optimal user experience. We partially responded to the aesthetics question by bringing the game's default graphics into the landing page.

For the judgment generator, we sent the link to about ten of our friends and acquaintances, who shared it with people in their orbit. The responses were unanimously positive. Within a few clicks, the generator demystifies the judging language by breaking it into three structural elements (intensifier, adjective, noun) and recombining those in often absurd and, at times, hilarious ways.

Overall, the interactive elements successfully encourage a wider audience reach, and our findings reflect a cohesive investigation and were well received at the final presentation. The project's sense of humor was mentioned and appreciated by audience members and classmates throughout the process, and it further contributes to inviting a general audience.

Future of the Project

As a practical matter, we will transfer our project in its current form from its independent, professional, and costly WordPress platform to the CUNY Commons, where it will be accessible to the public for a minimum of three years. The corpus will remain publicly available on GitHub for the foreseeable future unless something happens to the site itself, in which case we will migrate it to Quarto for future hosting.

Over the next few months, some additional findings relevant to the current corpora will be added to the site. This will include a closer look at the judging language's nouns and a comparison between the individual judges' use of language. Refining the integration of the games is also a goal for the next iteration. In the fall, when a new season of *GBBS* is scheduled to debut on

Netflix, selected findings will be used to pitch a piece of cultural criticism to relevant publications.

Eventually, our current findings will likely become integrated into the scope of a larger text analysis project. *Ob&Up* could mark the beginning of a larger investigation of the language of TV judging and/or screen-mediated language about food. One goal would be to expand the corpus to include additional *GBBS* seasons and consider the findings in aggregate. This would allow us to look at controversial aspects of the show, like the judging language in themed episodes (e.g., Mexican Week in Season 13) that have been rightfully criticized for cultural appropriation.

Since *GBBS* has several direct spin-offs and has inspired a slew of less directly related baking shows, creating a broader corpus (or several comparison corpora) would be the beginning of a more significant query into the language of food-related TV judgment, which would amplify the central goal of *Ob&Up*: to create consciousness around our increasingly lopsided (and reduced) linguistic relationship to multi-sensory experiences.