Sitcom Form and Function: Pacing and Production in a Collection of Thirty U.S. Series

Taylor Arnold[1], and Lauren Tilton[2]

[1] Data Science and Linguistics, University of Richmond, U.S.A.
[2] Rhetoric and Communication Studies, University of Richmond, U.S.A.

Abstract

Using an original corpus of thirty U.S. situational comedies (sitcoms) spanning over 4,500 episodes, we investigate how visual and aural pacing evolve over time and across modes of production. We find a clear trend toward faster pacing in visual editing, spoken dialogue, and textual density throughout the decades. While this shift correlates strongly with changes between multi-camera and single-camera setups, it is also shaped by the narrative goals of each series. For example, *Seinfeld* and *Frasier*, despite sharing visual and production similarities with other 1990s multi-camera sitcoms, feature markedly faster dialogue that reflects a narrative emphasis on wit and language. In contrast, single-camera series such as *Modern Family* and *The Office (US)* combine rapid dialogue with long takes and visual pauses that support physical and situational humor.

By combining large-scale computational analysis with close attention to aesthetic and narrative function, this study contributes to ongoing debates in television theory regarding the relationship between form and meaning. Whereas some scholars have emphasized the sitcom's formal conservatism and narrative stability, our findings reveal a more dynamic interaction between production technology, pacing, and storytelling strategy. Drawing from media-specific approaches and cultural theory, we argue that sitcom style emerges through a negotiation between material affordances and discursive intentions. This approach reframes how we understand the evolution of sitcom aesthetics and offers new empirical insight into the genre's formal diversity and cultural significance.

Keywords: computational television studies, multimodal analysis, situational comedies, distant viewing, digital humanities

1 Introduction

In the post-World War II era, television rapidly emerged as a dominant medium for both news and entertainment throughout the United States. By the late 1950s, a majority of U.S. households owned a television set [4]. The technology evolved dramatically over subsequent decades: transitioning from small black-and-white receivers with signals transmitted over airwaves to color television becoming standard by the late 1960s, followed by cable television in the 1980s, high-definition broadcasting in the 2000s, and most recently streaming services such as Netflix and Hulu. Despite these technological transformations, the medium has proven remarkably resilient in its fundamental structures. Even contemporary direct-to-stream productions typically conform to established temporal formats (approximately 22 or 44 minutes of content) and fit within familiar genre categories inherited from broadcast television.

Among television's most enduring formats, situational comedies, commonly known as sitcoms, have maintained consistent popularity from the 1950s to the present day. These comedic

Taylor Arnold, and Lauren Tilton. "Sitcom Form and Function: Pacing and Production in a Collection of Thirty U.S. Series." In: *Computational Humanities Research 2025*, ed. by Taylor Arnold, Margherita Fantoli, and Ruben Ros. Vol. 3. Anthology of Computers and the Humanities. 2025, 232–248. https://doi.org/10.63744/yHo626es4FhQ.

series employ a narrative structure centered on a core set of characters and locations, with common configurations revolving around family units (*Fresh Prince of Bel-Air*), friendship groups (*Living Single*), or workplace environments (*The Office*). Sitcoms traditionally feature an episodic structure wherein each episode functions independently and most plotlines reach resolution within a single installment. This format proved ideally suited to the Network-Era concept of a "least objectionable program," an idea from NBC executive Paul L. Klein which prioritized content designed to minimize viewer objections across the broadest possible audience. Series such as *The Donna Reed Show* and *The Dick Van Dyke Show*, with focuses on domestic situations and broadly appealing humor, exemplify this approach.

The proliferation of channels during the multi-channel transition era (mid-1980s to late 1990s) transformed rather than diminished the sitcom's cultural significance. This period witnessed an explosion of series targeted at increasingly specific demographic segments defined by gender, age, race, and geographic location. Subsequently, the shift toward what Jason Mittell identifies as "complex TV" has further evolved the sitcom format without displacing it [27]. Contemporary series including *Arrested Development* and *Community* have successfully integrated intricate serialized narratives and experimental production techniques into the traditional sitcom framework, demonstrating the format's continued capacity for innovation.

Despite sitcoms' seven-decade dominance in American entertainment, their formal qualities remain understudied due to multiple factors[19]. First, sitcoms have historically been dismissed as formulaic and stylistically uninteresting, particularly those produced during the Network Era. Second, and perhaps more significantly, the sheer scale of television presents formidable methodological challenges. A complete viewing of a single long-running series can require 50 to 100 hours or more; systematic analysis of even a modest collection of series demands thousands of hours of viewing time. These practical constraints have limited most television scholarship to selective sampling, plot summaries, or reliance on secondary sources.

The emergence of computational methods for analyzing audiovisual media offers a solution to these longstanding limitations. Recent advances in computer vision, speech processing, and machine learning now enable researchers to extract and analyze rich multimodal features across extensive corpora at scales previously impossible. In this paper, we demonstrate how computational techniques, combined with traditional close analysis and historical contextualization, can reveal patterns in the formal evolution of television sitcoms that have remained difficult to discern using conventional scholarly methods. By examining visual elements (shot duration, face presence), aural features (speech rates, turn-taking patterns), and their relationships across thirty series spanning seven decades, we uncover how sitcoms have adapted their formal strategies to changing production contexts, audience expectations, and cultural dynamics while maintaining their fundamental appeal as a television format.

2 Prior Work

Influenced by literary studies and art history, film studies has long embraced close analysis of form, style, and content in feature-length films [13]. Research has integrated quantitative measurements into film analysis, with scholars including David Bordwell [5], Gunars Civjans [37], Barbara Flueckiger [16], Daria Khitrova [3], Barry Salt [31], and Yuri Tsivian [36] pioneering computational approaches. This work primarily focused on shot duration analysis and textual analysis of subtitles and transcripts. While early studies used manual annotation, recent research incorporates automatically generated features, such as gender-based character presence detection [2].

Beyond simple measurements of shot length or face counts, researchers have shown interest in more complex features, though data annotation challenges have limited such investigations. Barry Salt's analysis of split edits across 33 films exemplifies this challenge [30]. He manually coded J-edits (audio from the next scene preceding the visual cut) and L-edits (audio continuing after

the visual cut) as binary features. Despite never achieving the capacity for large-scale analysis, Salt's *Statistical Style Analysis of Motion Pictures* provides detailed descriptions of features such as camera movement, angle, and audio characteristics that would prove highly insightful if studied at scale [31].

Television's influence on American political, social, and cultural life has been extensively documented [23; 24; 25; 26; 34]. Numerous series and episodes have contributed significantly to national discourse on issues ranging from feminism to civil rights to urban-rural divides [7; 15; 38]. However, unlike film studies, television scholarship has relied primarily on archival studio records, plot summaries, historic reviews, and other textual documents rather than analyzing the audiovisual content itself. This reliance on paratextual sources has shaped the kinds of questions that television scholars are able to pursue. Because these materials emphasize narrative, production history, and reception, they tend to foreground themes, character arcs, and industry context while leaving aside the formal features of the medium. As a result, elements such as pacing, shot duration, editing rhythm, and audiovisual density have often received less sustained attention. When these aspects are addressed, they are frequently discussed through anecdotal close readings, without a broader framework for comparison across series or time periods.

Although computational approaches to television are gaining attention, most existing work has focused narrowly on dialogue transcripts and narrative content. These studies often adapt methods from computational literary analysis and tend to overlook the visual and formal dimensions that are central to the medium [10]. As a result, features such as shot composition, editing pace, and audiovisual rhythm remain largely unexamined at scale. Jeremy Butler stands out as one of the few scholars applying computational analysis to television's visual components. His shot-length analysis of *Happy Days* (1974–1984) provides particularly insightful observations about editing style in sitcoms [8]. Similarly, Arnold, Berke, and Tilton examined how shot types signify character relationships in *Bewitched* (1964–1972) and *I Dream of Jeannie* (1965–1970) [1]. To our knowledge, no previous computational studies of narrative television have approached the scale of our current investigation, nor have any examined the interplay between visual and aural structures using computational methods.

3 Data

We have assembled an original dataset containing the audiovisual contents of episodes from thirty U.S. sitcoms. Data collection involved transcoding DVDs and Blu-ray discs purchased through our institution, utilizing the research exemption to DMCA §1201 under U.S. copyright law, which permits breaking digital rights management systems for academic research [12]. The thirty series were selected based on three criteria: commercial availability of DVDs or Blu-ray disks, representation in existing scholarship, and coverage of popular sitcom types from the mid-1950s to the present. While achieving a truly "random" sample of all U.S. sitcoms would be impossible, our collection includes many of the most historically significant series and provides a robust foundation for analyzing the medium's evolution.

Table 1 lists all shows in our corpus chronologically by premiere date. We included every available episode, though three shows are incomplete due to availability constraints. *The Donna Reed Show* has only its first five seasons commercially available. *Black-ish* released only its first season on DVD before transitioning exclusively to streaming platforms. *My Living Doll* (1964-1965) survives with only ten episodes from its single season. Despite commercial failure, we include it for historical importance and connections to other fantasy-based sitcoms in our collection. We also include *The Good Place*, which, though more fantasy-comedy than traditional sitcom, provides valuable comparison to early fantasy series (*Bewitched*, *I Dream of Jeannie*, and *My Living Doll*). Notably, its creator Michael Schur also produced three other sitcoms in our collection (*The Office (U.S.)*, *Parks and Recreation*, and *Brooklyn Nine-Nine*), which also places it as an

Series	Years	Camera	S#	E#	Dur.	Cast	Type	FPS
I Love Lucy	1951-1957	Multi	6	179	26.1	4	SD	24
Donna Reed Show	1958-1966	Single	†5	186	25.6	4	SD	30
Dick Van Dyke Show	1961-1966	Single	5	158	25.5	5	HD	24
My Living Doll	1964-1965	Single	1	†10	25.3	4	SD	30
Bewitched	1964-1972	Single	8	254	25.3	5	SD	30
I Dream of Jeannie	1965-1970	Single	5	138	25.0	4	SD	24
Mary Tyler Moore Show	1970-1977	Multi	6	144	25.5	4	SD	24
All in the Family	1971-1979	Multi	9	202	25.1	4	SD	30
Sanford and Son	1972-1977	Multi	6	135	24.9	2	SD	30
Good Times	1974-1979	Multi	6	133	25.2	4	SD	30
Cheers	1982-1993	Multi	12	270	24.0	8	SD	24
Seinfeld	1989-1998	Multi	9	165	22.8	4	SD	24
Fresh Prince	1990-1996	Multi	6	146	22.6	6	SD	30
Living Single	1993-1998	Multi	5	118	22.4	6	SD	30
Frasier	1993-2004	Multi	11	256	22.1	5	SD	30
Friends	1994-2004	Multi	10	228	23.4	6	SD	30
Everyb. Loves Raymond	1996-2005	Multi	9	207	22.4	6	SD	24
That '70s Show	1998-2006	Multi	8	200	21.9	9	SD	30
Arrested Development	2003-2006	Single	3	52	22.0	9	SD	24
The Office (US)	2005-2013	Single	9	185	22.1	14	SD	24
How I Met Your Mother	2005-2014	Multi	9	205	21.5	5	SD	24
30 Rock	2006-2013	Single	7	133	21.3	8	HD	24
The Big Bang Theory	2007-2019	Multi	12	279	20.4	7	SD	24
Community	2009-2014	Single	6	98	21.3	9	SD	24
Parks and Recreation	2009-2015	Single	8	122	21.5	8	SD	24
Modern Family	2009-2020	Single	11	249	21.6	11	SD	24
Brooklyn 99	2013-2021	Single	8	153	21.6	9	HD	24
Black-ish	2014-2022	Single	†1	24	21.5	9	SD	24
The Good Place	2016-2020	Single	4	48	24.5	6	HD	24

Table 1: Summary statistics for the situational comedies in our dataset, sorted by the first year of distribution. Indicates the camera setup type, total number of seasons (S#), the total number of episodes (E#), the median episode duration in minuates (Dur.), the main cast size (Cast), the screen resolution of our data, and the median frames per second. Further details are given in the text. When available, we include data from the original run of each show. Counts with a † indicate that our set is only a subset of the full show due to data availability.

interesting case-study to understand the limits of using a sitcom format to tell complex narratives.

Before jumping into further computational analysis, the metadata in Table 1 already reveals several compelling patterns. Episode duration has steadily decreased from 25-26 minutes in the 1950s-1960s to 20-21 minutes by the mid-2010s, reflecting changing commercial practices and audience expectations. Main cast sizes have generally increased over time, with recent shows featuring notably large ensembles: *The Office (U.S.)* (14 members) and *Modern Family* (11 members). Technical specifications, including HD availability and frame rates, largely depend on original production methods. Shows shot on tape (common in the 1980s-1990s) exist only in standard definition, while DVD sets predating Blu-ray dominance have rarely been reissued except for the most popular series. Fortunately, as our results demonstrate, video resolution does not significantly

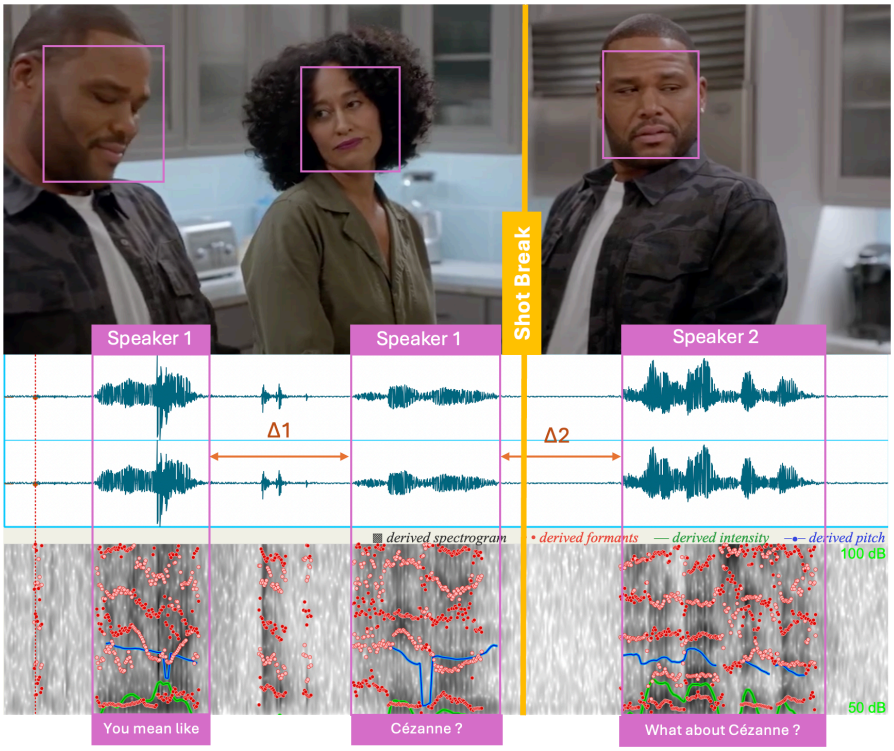

Figure 1: Example of shot detection, face detection, speech detection, speaker diarization, and transcription from a clip of an episode of *Black-ish* (Season 1, Ep. 9; 06:00-06:04).

.

affect our analytical outcomes.

Perhaps the most significant metadata distinction involves camera setup, which profoundly impacts writing and production approaches [9]. Single-camera setups record from one camera position at a time, requiring repositioning of equipment between shots. This approach offers maximum control over visual aesthetics and enables location shooting, producing a more cinematic style. Multi-camera setups operate three or four cameras simultaneously, capturing different angles and shot scales in real time. While requiring more equipment, this method reduces production time and costs significantly. The format's theatrical quality stems from actors performing entire scenes continuously, making it ideal for live audiences. *I Love Lucy* pioneered the multi-camera approach for scripted comedy [32], establishing it as the dominant sitcom format from the 1970s through early 2000s. Today's productions utilize both formats, with the choice significantly influencing each show's visual and narrative style.

4 Methods

4.1 Algorithms

After standardizing our corpus to MP4 format, we applied audiovisual algorithms to every episode. Figure 1 illustrates the complete pipeline of our analytical process.

We began with shot boundary detection, ultimately selecting the TransnetV2 algorithm after extensive testing confirmed its reliability across our diverse collection [33]. The algorithm generates frame-specific predictions with associated probability scores. After shot detection, we extracted middle frames and applied face detection using the `buffalo-large` algorithm from the InsightFace module with a 0.7 confidence threshold [17]. We chose middle-frame extraction to avoid

Algorithm	Class. Rate (Overall)	Class. Rate (Series Range)	Hardest Series
Shot Boundary Detection	99.3%	90%–100%	*That 70s Show*
Number of Faces	98.3%	95%–100%	*The Good Place* *Living Single*
Number of Speakers	93.0%	80%–100%	*All in the Family*
Speaker Gap	95.2%	85%–100%	*30 Rock* *Sanford and Son* *Good Times*
Transcription	99.5%	97.2%–99.7%	*Good Times*

Table 2: Classification rates of algorithms according to hand-labeled data consisting of 20 samples from each of the 30 series in the corpus. Classification rates are word-level error rates for the transcription task and binary error rates for the other tasks.

biasing results toward longer shots, which would naturally accumulate more face detections. To eliminate spurious background detections, our final counts include only faces comprising at least 70% of the width of the largest face in each frame.

For audio analysis, we extracted MP3 files from each video to generate aural features. The Whisper `large-v3` model provided time-stamped transcriptions, with English specified as the target language [29]. Results include predicted words, word-level timestamps, and confidence scores. We then applied PyAnnote's speaker diarization model [6]. This model generates utterance timestamps, confidence scores, and categorical codes linking utterances from the same speaker. While an open-source variant exists, the commercial API demonstrated significantly superior accuracy. Processing over 1,800 hours of material through the advanced model cost $272.

4.2 Evaluation

While our chosen models demonstrate high accuracy on their original training data, our corpus presents unique challenges absent from typical benchmarks. Our collection includes black-and-white footage, low-resolution images compared to HD training sets, and significantly more background noise than standard speaker detection datasets. These differences necessitate additional validation to ensure algorithm reliability on our source material.

We created an evaluation dataset by randomly selecting 20 detected shots from each of our 30 series, generating short video segments for each shot. The authors annotated: (1) shot detection accuracy, (2) foreground face count in the middle frame, and (3) unique speaker count within each shot. To evaluate audio algorithms, we additionally selected 20 detected utterances per series. Authors assessed word error rates in transcriptions and verified speech detection accuracy within ±50 milliseconds.

Table 2 presents algorithm error rates compared to our hand-labeled ground truth. The algorithms all had acceptable accuracy rates, with some algorithms performing better than others. The shot boundary detection and transcription tasks both had classification rates of over 99% over the whole corpus. The face detection algorithm and speaker gap detection were also fairly accurate, with values of 98.3% and 95.2% respectively. Detecting the number of speakers detection proved to be the most difficult task, with an accuracy rate of 93%. The majority of these errors were the result of laughter or other sound effects being classified as an additional speaker. Looking at the range of error rates across all series, as well as those series with the worst classification rates, shows that for no series are the classification rates noticiably worse than the average rates.

4.3 Statistical Summaries

Following established practice in shot duration research, which has documented heavy-tailed distributions with substantial outliers (often approximating log-normal distributions) we employ robust statistical measures throughout our analysis. We calculate median shot length (MSL) as our primary metric, consistent with film studies conventions. For other counts and durations, we use 10% symmetrically trimmed means to minimize outlier influences. To characterize the variability in these metrics, we report median absolute deviation (MAD) with normal correction for the MSL values and for trimmed means, we calculate trimmed standard deviations using the same 10% threshold.[1]

Our results focus on series-level summary statistics. While we include variability measures, we avoid statistical inference at the series level since, for all but three shows, we possess complete populations rather than samples. Beyond individual series, we deliberately avoid aggregated statistics across all 30 shows, recognizing that our dataset does not represent anything close to a random sample from all U.S. sitcoms. The only cross-series statistics we present are correlations examining relationships between aural, visual, and textual elements across our full collection, providing insight into how these components interact within the medium.

5 Results

The main results are in Tables 3–6. In this section, we give an overview of what data is represented in these tables and a summary of the patterns and outliers represented within them. Implications of these findings relative to existing television scholarship are further explored in the following section.

To examine the relationship between shot composition and editing rhythm, we analyzed median shot length (MSL) across different categories of visual character presence within shots and the number of speakers heard during the shot. Table 3 presents the resulting MSL values. These explore pacing decisions based on visual and aural character quantities per shot. Results are arranged by ascending MSL for comparison. To measure the variability of these measurements, Table 6 in the appendix provides the normally-adjusted median absolute deviation scores. The relationships shown in Table 3 reveal several distinct patterns between the temporal rhythm of the visual material as a function of shot length. Following other stylometric analyses of film and television, we see a general trend toward increasing shot lengths over time [8; 31]. This is not entirely a deterministic relationship, as we see, for example, several 1970s series with slower shot pacing than all of those from our set in the 1950s and 1960s. Examining the MSL scores by the number of faces and speakers allows us to explore this relationship in greater depth.

Unsurprisingly, the MSL increases when there are more speakers in a shot. We see that this pattern is particularly resilient. In every single series and for every category of the number of faces, the MSL increases between 1 and 2 speakers. These gaps can be relatively small. On the low end we have a 15% increase (0.3 seconds) for the difference between one and two speakers with one face in *Brooklyn 99*. In *The Donna Reed Show*, we see a 250% increase (5.6 seconds) in the MSL of shots with two faces when comparing the difference between one and two speakers. There is also an increase in MSL for every number of faces when comparing the difference between 2 and 3 speakers, with the sole exception being shots with one face in *The Good Place*, where 2.7 seconds (1 face and two speakers) decreases to 2.5 seconds (1 face and three speakers). Modern single-camera series such as *Brooklyn 99* and *Community* again show only modest increases of a few hundred milliseconds. On the high end, a few gaps are particularly large, though these correspond

[1] MAD is defined as the median value of the absolute difference of each value from the median of the sample. The normal correction multiplies this by 1.4826, a theoretically derived constant ensuring convergence to standard deviation for Gaussian distributions.

Series	MSL	1 Speaker			2 Speakers			3+ Speakers		
		F1	F2	F3	F1	F2	F3	F1	F2	F3
Brooklyn 99	2.0	2.0	2.0	1.9	2.3	2.5	2.4	2.4	2.9	2.9
Kim's Convenience	2.1	2.1	2.3	2.2	2.3	3.1	2.8	2.6	4.0	3.7
30 Rock	2.1	2.1	2.1	2.0	2.5	2.9	2.4	2.7	3.4	2.9
Fresh Off The Boat	2.1	2.3	2.3	2.2	2.6	3.0	2.8	2.6	3.7	3.0
Community	2.2	2.2	2.0	1.9	2.5	2.8	2.6	2.7	3.6	3.4
Black-ish	2.2	2.1	2.1	2.0	2.4	2.6	2.5	2.7	3.4	2.7
Parks and Recreation	2.3	2.3	2.1	2.0	2.6	2.9	2.7	3.4	3.7	3.8
Arrested Development	2.3	2.1	2.3	2.2	2.8	3.5	3.4	3.5	4.6	4.0
The Good Place	2.3	2.3	2.4	2.3	2.7	3.0	2.8	2.5	3.2	3.0
How I Met Your Mother	2.3	2.4	2.3	2.1	2.7	3.3	3.0	3.4	4.5	4.0
The Big Bang Theory	2.6	3.0	3.0	2.7	3.0	3.8	3.5	3.0	4.3	4.2
The Office (US)	2.7	2.6	2.5	2.4	3.1	3.6	3.2	4.8	5.6	4.9
Seinfeld	2.8	2.8	3.0	3.1	2.8	4.0	4.4	4.0	6.0	6.9
Friends	2.8	2.9	3.0	2.9	3.2	4.3	4.2	4.4	5.5	5.1
Modern Family	2.9	2.6	2.6	2.4	3.1	3.9	3.5	5.2	5.8	5.9
I Dream of Jeannie	3.0	2.6	3.1	3.7	3.8	8.6	7.8	14.8	20.7	17.0
Cheers	3.2	2.8	3.0	2.9	3.9	5.3	5.0	5.9	7.9	7.8
Bewitched	3.3	3.2	3.8	4.4	4.5	8.5	8.7	9.1	14.5	14.5
Frasier	3.3	3.0	3.4	3.4	3.5	4.9	4.8	4.9	6.7	6.8
Everyb. Loves Raymond	3.4	2.9	3.3	3.2	4.6	5.6	4.6	6.5	8.5	7.2
That '70s Show	3.8	4.0	4.2	4.3	4.4	6.2	5.8	7.6	8.6	8.5
Mary Tyler Moore Show	3.8	3.4	3.9	3.9	4.0	6.6	5.7	6.7	9.9	9.2
Living Single	3.8	4.0	4.4	3.9	4.6	6.9	5.7	6.8	9.3	8.3
My Living Doll	3.9	3.4	3.8	3.1	5.6	10.3	7.1	7.8	19.0	14.4
Donna Reed Show	3.9	3.3	3.8	4.3	4.2	9.4	9.4	15.4	20.3	19.8
Fresh Prince	4.0	3.7	4.5	4.0	4.6	7.6	7.0	9.1	10.5	9.6
Dick Van Dyke Show	4.1	3.1	3.2	4.3	4.4	6.1	6.7	10.0	10.3	11.0
I Love Lucy	4.4	3.0	3.6	4.6	4.1	5.8	6.1	6.3	9.6	10.8
Good Times	4.5	4.1	4.4	4.0	4.6	7.1	5.9	7.5	10.0	10.3
Sanford and Son	5.1	4.3	5.2	4.8	4.8	8.8	7.0	9.0	12.7	11.3
All in the Family	5.3	4.6	4.9	4.8	5.3	8.8	7.5	9.7	12.8	13.1

Table 3: Summary of the median shot length (MSL) in seconds by the number of faces present in the shot. The first column gives the overall MSL. The next three columns give the MSL for shots with one detect speaker according to the number of faces: F1 is one one face, F2 is two faces, and F3 is three or more faces. The next three columns give the same breakdown of MSL for shots with two speakers and the last three columns give the breakdown of MSL for shots with three or more speakers. The corresponding median absolute deviations are give in the appendix. The results are ordered by the overall MSL.

to relatively rare shot types, such as three speakers with only one face.

The MSL of shots in which the number of speakers is equal to the number of faces present in the middle frame (or both are greater than three) reveals another consistent pattern. Across all of the series, the MSL of one speaker and one face is less than or equal to the MSL for two speakers and two faces, which is itself less than or equal to the MSL of three or more speakers and three or more faces. While these all show a general increase, the variability in the differences between these

Series	MSL	1 Speaker			2 Speakers			3+ Speakers		
		F1	F2	F3	F1	F2	F3	F1	F2	F3
Brooklyn 99	2.0	36	11	4	22	7	2	2	1	0
Kim's Convenience	2.1	39	6	1	24	5	1	2	1	0
30 Rock	2.1	48	7	2	19	3	1	2	1	0
Fresh Off The Boat	2.1	44	9	4	14	4	1	1	1	0
Community	2.2	33	10	4	17	6	2	3	2	1
Black-ish	2.2	36	10	4	22	7	2	3	2	1
Parks and Recreation	2.3	41	11	5	18	6	2	1	1	0
Arrested Development	2.3	40	7	2	20	4	1	4	1	0
The Good Place	2.3	43	13	6	14	5	2	1	1	0
How I Met Your Mother	2.3	35	17	7	12	7	2	1	1	1
The Big Bang Theory	2.6	37	13	4	12	6	2	1	1	0
The Office (US)	2.7	35	8	3	18	6	2	3	2	1
Seinfeld	2.8	31	10	3	21	9	2	2	2	1
Friends	2.8	38	12	4	14	7	2	2	2	1
Modern Family	2.9	28	10	3	23	11	3	3	3	1
I Dream of Jeannie	3.0	28	7	2	11	7	2	2	3	2
Cheers	3.2	28	15	7	14	9	4	2	2	1
Bewitched	3.3	35	8	2	12	7	2	1	2	1
Frasier	3.3	34	10	4	20	9	3	3	2	1
Everyb. Loves Raymond	3.4	29	9	3	19	8	3	3	2	1
That '70s Show	3.8	37	14	7	9	6	3	1	1	1
Mary Tyler Moore Show	3.8	28	8	2	21	9	2	4	4	2
Living Single	3.8	34	12	5	11	8	3	2	2	1
My Living Doll	3.9	31	9	2	11	11	2	1	3	2
Donna Reed Show	3.9	28	7	2	16	9	2	3	4	2
Fresh Prince	4.0	30	12	5	10	9	3	2	3	1
Dick Van Dyke Show	4.1	25	10	3	18	12	3	4	5	3
I Love Lucy	4.4	17	7	3	20	11	4	7	7	4
Good Times	4.5	25	9	4	17	11	4	4	4	3
Sanford and Son	5.1	25	7	2	19	12	3	3	4	2
All in the Family	5.3	29	5	1	23	10	2	5	4	2

Table 4: The distribution of shots by the number of speakers, with the shots with no speakers removed. The first column gives the overall MSL in seconds, which was used to order the results and correspond with Table 3. All other results are given as percentages. The three columns under 'Speaker 1' give the distribution of shots with one speaker and the following number of faces: F1 for one face, F2 for two faces, and F3 for three or more faces. The next three columns give the same for shots with two speakers and the last trhee columns give the distribution for shots with three or more speakers.

three MSL values appears to be a strong differentiator in style across our corpus. For example, *Cheers* and *Bewitched* have similar overall MSL values of 3.2 and 3.3 seconds. However, for *Cheers* the MSL values for matching face and speaker scores are 2.8, 5.3, and 7.8 seconds. For *Bewitched*, these values increase substantially to 3.2, 8.5, and 14.5 seconds. In general, there is a strong temporal pattern in values of two speakers with two faces and three or more speakers with three or more faces. All of the shows that premiered before 1980 have an MSL for two faces with

Series	Words Per Minute		Turn Duration	Gap Size
	Speaking	Overall		
That '70s Show	229 (17)	120 (8)	3.42 (2.76)	1.32 (1.22)
I Love Lucy	242 (21)	123 (18)	1.81 (1.53)	0.31 (0.50)
My Living Doll	240 (15)	130 (28)	2.66 (2.19)	0.49 (0.50)
Bewitched	240 (18)	131 (13)	2.61 (2.31)	0.58 (0.63)
Friends	240 (11)	136 (10)	2.36 (2.09)	0.71 (0.86)
I Dream of Jeannie	262 (17)	139 (14)	2.12 (1.81)	0.40 (0.51)
Cheers	237 (14)	144 (10)	2.63 (2.36)	0.65 (0.79)
Living Single	241 (11)	144 (12)	2.89 (2.57)	0.78 (0.90)
Everyb. Loves Raymond	238 (15)	144 (16)	2.51 (2.00)	0.77 (0.98)
Fresh Prince	250 (17)	149 (8)	2.85 (2.40)	0.81 (0.93)
The Big Bang Theory	255 (11)	150 (12)	2.88 (2.15)	0.91 (0.95)
Donna Reed Show	251 (14)	150 (17)	2.31 (1.99)	0.46 (0.52)
Sanford and Son	259 (19)	150 (19)	2.55 (2.06)	0.42 (0.57)
The Office (US)	245 (18)	153 (10)	2.35 (2.13)	0.39 (0.53)
How I Met Your Mother	251 (16)	156 (13)	2.90 (2.55)	0.48 (0.59)
Good Times	239 (11)	157 (8)	2.55 (2.21)	0.40 (0.60)
Fresh Off The Boat	242 (13)	157 (11)	2.98 (2.36)	0.43 (0.58)
Community	249 (21)	160 (13)	2.28 (2.22)	0.29 (0.45)
All in the Family	237 (16)	161 (14)	2.79 (2.19)	0.39 (0.64)
30 Rock	248 (15)	162 (12)	2.99 (2.63)	0.28 (0.44)
Seinfeld	257 (12)	164 (8)	2.04 (1.70)	0.49 (0.61)
Mary Tyler Moore Show	246 (12)	164 (10)	2.10 (1.84)	0.39 (0.50)
Dick Van Dyke Show	268 (21)	164 (25)	1.98 (1.58)	0.32 (0.41)
The Good Place	226 (5)	165 (7)	4.22 (3.48)	0.31 (0.45)
Frasier	260 (16)	166 (11)	2.55 (2.17)	0.47 (0.65)
Parks and Recreation	257 (23)	172 (9)	3.10 (2.54)	0.31 (0.40)
Modern Family	267 (11)	180 (8)	2.50 (1.99)	0.21 (0.33)
Black-ish	243 (12)	180 (10)	2.21 (2.12)	0.14 (0.25)
Arrested Development	259 (10)	186 (14)	2.25 (2.07)	0.24 (0.32)
Brooklyn 99	265 (15)	187 (8)	2.50 (2.14)	0.12 (0.22)

Table 5: Summary of speech within the corpus of U.S. sitcoms. The first two columns of results show the words per minute, with the moments of speech and the total show as denominators, respectively. The second two columns provide the length of a speech turn taken by a speaker and the duration between speakers. Both of these are measured in seconds. All results are given as the trimmed means and trimmed standard deviations (10%).

two speakers of 5.8 seconds or greater and an MSL for the corresponding case with three or more speakers and faces of 10.3 seconds or greater. All of the shows from the 1990s onward, regardless of the camera type, have an MSL of less than 3.9 seconds for the two character case and less than 5.9 seconds for the three or more character case. For shows from the 1980s and 1990s, these two values tend to fall somewhere between these ranges.

For a fixed number of speakers, the relationship between shot length and the number of faces present is less stable across the corpus. One clear pattern that emerges is that in the case of two speakers, the MSL with one face present is less than the MSL with two faces present for every series in the corpus. For one speaker, there seems to be little overall difference between the shot

length and the number of faces. For two and three speakers, other than the one relationship already mentioned, no clear pattern emerges. This is likely in part due to the different ways that a single face can be present. For example, *Brooklyn 99* features a large number of panning shots, so only having one face in the central frame of the shot may still correspond to multiple characters being visible at some point during the shot. In the case of the multi-camera shows, two faces corresponds to two different, but both popular, shot types: the over-the-shoulder shot and the two-shot. Additional work could help reveal more patterns relative to these more granular shot types. Another influence is the lower number of examples of certain combinations of speakers and faces. We can look at the distribution of these shots, in addition to their length, to further understand the style of the sitcoms in our collection.

There is a general pattern toward more shots with one speaker over time. In Table 4, we show the percentage of shots from each series that have a given number of faces and speakers. For every series other than *I Love Lucy*, the most common shot type has a single speaker and a single face. Other combinations of one to two speakers and one to two faces are the next most common for all of the series. The differences across shows can be striking. *I Love Lucy* has only 25% of the shots having a single speaker compared to the 70% rate for *Fresh Off the Boat*. These are also extremes in terms of the percentage of shots with three or more speakers, with 35% and 3%, respectively. The increase in the percentage of one speaker seems to be less closely related to shot type compared to the number of faces. Several multi-camera shows from the 2000s such as *That '70s Show* and *How I Met Your Mother* have higher rates of a single speaker than about half of the modern single-camera shows. The post-2000 single-camera sitcoms reveal a unique pattern: they are the only shows that have more than 66% of their shots with only a single face. Looking back at Table 3, these also tend to be the series with the lowest MSL values overall, with the mockumentaries *The Office (US)* and *Parks and Recreation* being slightly slower. In general, we see that the increasing pace of series over time is a two-fold process, resulting from both having more shots with a single character in the foreground (in particular, for the single-camera shows) and having faster shots even when two characters are present.

The rate of speech in a sitcom series is both a confounding factor for the relationship between MSL and the number of speakers, as well as an interesting feature in its own right for the analysis of pacing and production. Speech delivery patterns within the U.S. sitcom corpus exhibit distinctive characteristics that reflect both genre conventions and performance practices in television comedy. Table 5 characterizes the verbal landscape of these programs through multiple complementary measures of dialogue intensity and pacing. The metrics encompass both the density of spoken content, calculated against active speech time and total program duration, and the temporal structure of conversational exchanges. This includes the duration of individual speaking turns and the intervals that separate them. Using robust statistical measures that minimize the influence of extreme values, these findings capture the central tendencies and variation in how sitcom dialogue unfolds.

The overall words per minute spoken on each series shows strong stylistic differences across their textual densities. We see that the series' overall words per minute range from 120 to 187. This range corresponds to natural rates of speech that have been observed in spoken English [20]. In general, the series with the highest textual density are modern single-camera sitcoms. However, there is substantial variation that appears to be related to the specific narrative aims of each series. The witty single-camera sitcoms *Seinfeld* and *Frasier* have higher words per minute than the modern single-camera series *Community*, *30 Rock*, and *The Office (US)*. *The Dick Van Dyke Show* and *The Mary Tyler Moore Show* also have higher average words per minute than the other Network Era sitcoms. It is possible that this is due to the inclusion of music and simulated news broadcasts, respectively, both of which tend to have a higher density of speech [21]. The rate of speech when characters are speaking is more compressed and less variable. With the exception of *That '70s*

Show, the rate is between 240 and 267 words per minute, with a modest increase correlated with the overall density of speech in each series. The most interesting aspects of the average turn duration and gaps between speech are the outliers, which also fall within reported corpus-based data of natural speech [11; 22; 35]. We again see that speech on *That '70s Show* is unique, being delivered in long slow chunks, with the second-largest turn length and largest gap size. *The Good Place* has the largest turn length, which may be due to its non-episodic structure and the need for longer turns to move the plot forward over each episode. Otherwise, the turns all seem on average to be around 2 to 3 seconds long with no clear production or temporal pattern. Additional annotations are likely needed to untangle this complex relationship.

The statistical results are one way to under production, form, and narrative of television series. While television studies has generally paid less attention to close analysis of specific formal elements, there are several debates about the relationship between production style and narrative structure that our statistics offer a len into. We now turn to how the numerical results can make contributions to TV scholarship.

6 Connections to Television Scholarship

Extensive scholarship has argued for the role of television sitcoms as a static and conservative genre formation. Production elements such as camera style, resolution, or color are not seen as central to meaning-making. In contrast, sitcoms are categorized as a relatively stable genre designed to fulfill social, economic, and narrative functions that persist across technological or aesthetic changes. A particularly reductive version of this approach is exemplified by Marshall McLuhan's provocation that "the medium is the message" [26]. However, even some of McLuhan's strongest critics continue to highlight the stability of television forms and the fact that while culture strongly influences television forms, the cultural impact of the visual forms themselves generally remain limited. Raymond Williams, for example, has argued that technology, of which television is one example, is deeply influenced by cultural phenomena, and these cultural features influence the modes of production within television [41]. At the same time, Williams suggests that the most important aspects of the television medium are fixed features such as linearity, episodic structure, and flow. In his view, it would take fundamentally new modes of production (possibly anticipating the internet and streaming) before real structural change could occur. Similarly, Jason Mittell's work on television genre [28] and John Ellis' work on repetition tend to highlight the importance of static television features over the dynamic and changing modes of production [14]. Changes in production format, in other words, are stylistic evolutions that do not significantly alter the genre's core cultural and economic functions.

In contrast to these theories, other media and television scholars have emphasized the influence that modes of production have had on the ways and possibilities of stories that are told in sitcoms. For example, John Caldwell's *Televisuality* argues, with a particular focus on the post-Network Era, that production choices such as camera setup, cast size, and canned laughter directly shape audience perception, meaning, and affect. Caldwell argues that aesthetics and ideology are deeply entwined and informed by one another. Similarly, Lynn Spigel has argued that the importance of color television can be linked to notions of modernity and family ideology in the 1970s [34], whereas Kristen Warner has shown how the interaction of high production values has influenced the depiction of race and class on modern television series [39; 40]. With a focus on early 1930s television in Britain, which featured a wide variety of different production forms, Jason Jacobs provides a detailed description of how the interplay between form and function operates both visually and aurally [18].

Our study provides novel insights into the different theories regarding the relationship between form and function in U.S. sitcoms. While certain features such as overall episode length and dialogue dynamics such as turn length and speaker gaps show consistency over time, the formal

properties of sitcoms are far from static. Rather, we see a clear and noticeable trend toward increased speed and complexity, with shorter cuts, more dialogue, and a focus on fast sequences centered on shots of individual characters with audio constrained to individual speakers within a single shot. These findings point to a fluid definition of the sitcom genre that can adapt to technological innovations, as opposed to the conservative formations of genre offered by McLuhan and Ellis. The correlation between changes in these metrics and the disappearance and return of the single-camera setup in our dataset provides further evidence that these temporal changes are directly mediated through technological influences. Individual variations in our dataset—such as the lower speech density used to create awkward humor in *Modern Family* or the fast witty dialogue in the standard-definition multi-camera series *Seinfeld*—suggest that technological changes afford certain styles and meanings but do not deterministically dictate them. The output is filtered through narrative goals, cultural conventions, and audience expectations. This aligns closely with the qualitative conclusions drawn by Caldwell and Jacobs in their respective periods of interest. Further research will, hopefully, continue to provide similar nuanced perspectives on existing television scholarship regarding these and other phenomena.

7 Conclusions and Future Directions

We have seen a number of strong general patterns and interesting outliers, as well as developed connections between these results and existing television scholarship. As an overarching pattern, we have shown a clear trend toward faster pacing across multiple dimensions of televisual style. Our findings reveal consistent increases in editing speed, dialogue density, and textual compression from the 1950s through the 2010s, with median shot lengths decreasing substantially and words per minute increasing across the corpus. While these changes correlate strongly with the technological shift from multi-camera to single-camera production, our analysis demonstrates that pacing decisions are not merely determined by production constraints but are strategically deployed to serve specific narrative and comedic functions. Series such as *Seinfeld* and *Frasier* exemplify how creators can accelerate dialogue within traditional multi-camera frameworks to emphasize verbal wit, while single-camera series can achieve comedic effects through the calculated juxtaposition of rapid speech with extended visual holds. These patterns reveal that sitcom style emerges from a complex negotiation between material production affordances and aesthetic intentions, challenging previous characterizations of the genre as formally conservative and demonstrating instead a dynamic relationship between technological capabilities, narrative strategy, and comedic expression.

At a larger level, our contribution provides new evidence regarding the evolution of editing style and dialogue pacing in television, revealing how these stylistic choices have adapted to changing viewer expectations and technological capabilities. As Jeremy Butler observes, we must "tap into the production culture of a particular time in order to understand stylistic conventions" [8, p. 10]. By situating our computational findings within television production's historical and industrial contexts, we can begin to develop a nuanced understanding of how editing decisions reflect both creative intentions and cognitive principles of human perception.

Future directions for this research include a new materialist approach at the scene level, as well as a need to expand our corpus to include dramatic series and international productions, developing more sophisticated algorithms for detecting complex audiovisual patterns, and investigating how streaming platforms' binge-watching affordances may be reshaping fundamental timing conventions. Additionally, we plan to explore how our findings might inform contemporary production practices and contribute to media literacy education. Ultimately, this work demonstrates how computational methods can reveal the intricate formal systems through which television constructs meaning, exercises influence, and maintains its position as a dominant cultural force.

References

[1] Arnold, Taylor and Tilton, Lauren. *Distant Viewing: Computational Exploration of Digital Images*. MIT Press, 2023.

[2] Bamman, David, Samberg, Rachael, So, Richard Jean, and Zhou, Naitian. "Measuring diversity in Hollywood through the large-scale computational analysis of film". In: *Proceedings of the National Academy of Sciences 121*, no. 46 (2024), e2409770121. DOI: 10.1073/pnas.2409770121.

[3] Baxter, Mike, Khitrova, Daria, and Tsivian, Yuri. "Exploring cutting structure in film, with applications to the films of DW Griffith, Mack Sennett, and Charlie Chaplin". In: *Digital Scholarship in the Humanities 32*, no. 1 (2017), pp. 1–16.

[4] Boddy, William. *Fifties Television: The Industry and Its Critics*. Urbana: University of Illinois Press, 1990.

[5] Bordwell, David. *The Way Hollywood Tells It: Story and Style in Modern Movies*. 1st ed. University of California Press, 2006. (Visited on 04/22/2025).

[6] Bredin, Hervé, Yin, Ruiqing, Coria, Juan Manuel, Gelly, Gregory, Korshunov, Pavel, Lavechin, Marvin, Fustes, Diego, Titeux, Hadrien, Bouaziz, Wassim, and Gill, Marie-Philippe. "pyannote.audio: neural building blocks for speaker diarization". In: *ICASSP 2020, IEEE International Conference on Acoustics, Speech, and Signal Processing*. Barcelona, Spain, 2020.

[7] Charlotte Brunsdon and Lynn Spigel, edited by. *Feminist Television Reader*. Second edition, 2010. Oxford University Press, 2007.

[8] Butler, Jeremy. "Statistical analysis of television style: What can numbers tell us about TV editing?" In: *Cinema Journal* (2014), pp. 25–44.

[9] Butler, Jeremy G. *Television Style*. 1st. New York: Routledge, 2010.

[10] Chang, Kent K., Ho, Anna, and Bamman, David. "Subversive Characters and Stereotyping Readers: Characterizing Queer Relationalities with Dialogue-Based Relation Extraction". In: *Proceedings of the Computational Humanities Research Conference*, ed. by Wouter Haverals, Marijn Koolen, and Laure Thompson. Vol. 3834. CEUR Workshop Proceedings. 2024, pp. 917–937.

[11] Corps, Ruth E., Knudsen, Birgit, and Meyer, Antje S. "Overrated gaps: Inter-speaker gaps provide limited information about the timing of turns in conversation". In: *Cognition 223* (2022), p. 105037. ISSN: 0010-0277.

[12] Dombrowski, Quinn and Tilton, Lauren. "Access and Advocacy: Text & Data Mining and DMCA § 1201". In: *Digital Studies/Le champ numérique* , no. Special DSCN Collection# 9 (2024).

[13] Eisenstein, Sergei. *Film form: Essays in film theory*. Harcourt, Brace, 1949.

[14] Ellis, John. *Visible Fictions: Cinema, Television, Video*. Routledge, 1992.

[15] Emami, Gazelle. "How to Make It as a Black Sitcom: Be Careful How You Talk About Race". In: *The Huffington Post* (2014).

[16] Flueckiger, Barbara. "A digital humanities approach to film colors". In: *Moving Image: The Journal of the Association of Moving Image Archivists 17*, no. 2 (2017), pp. 71–94.

[17] Hast, Anders. "Age-Invariant Face Recognition Using Face Feature Vectors and Embedded Prototype Subspace Classifiers". In: *Advanced Concepts for Intelligent Vision Systems*, ed. by Jaques Blanc-Talon, Patrice Delmas, Wilfried Philips, and Paul Scheunders. Cham: Springer Nature Switzerland, 2023, pp. 88–99.

[18] Jacobs, Jason. *The intimate screen: Early British television drama*. Oxford University Press, 2000.

[19] Jullier, Laurent and Laborde, Barbara. *L'Analyse des Séries*. Armand Colin, 2024.

[20] Kowal, Sabine, Wiese, Richard, and O'Connell, Daniel C. "The Use of Time in Storytelling". In: *Language and Speech 26*, no. 4 (1983), pp. 377–392.

[21] Laver, John. *Principles of phonetics*. Cambridge university press, 1994.

[22] Levinson, Stephen C. "Turn-taking in Human Communication – Origins and Implications for Language Processing". In: *Trends in Cognitive Sciences 20*, no. 1 (2016), pp. 6–14. ISSN: 1364-6613.

[23] Lotz, Amanda D. *Redesigning women: Television after the network era*. University of Illinois Press, 2010.

[24] Lotz, Amanda D. "The television will be revolutionized". In: *The Television Will Be Revolutionized, Second Edition*. New York University Press, 2014.

[25] McLuhan, Marshall. *The Gutenberg Galaxy*. Toronto: University of Toronto Press, 1962.

[26] McLuhan, Marshall. *Understanding Media: The Extensions of Man*. Publisher, 1964.

[27] Mittell, Jason. *Complex TV: The Poetics of Contemporary Television Storytelling*. New York: New York University Press, 2015.

[28] Mittell, Jason. *Genre and Television: From Cop Shows to Cartoons in American Culture*. New York: Routledge, 2004.

[29] Radford, Alec, Kim, Jong Wook, Xu, Tao, Brockman, Greg, McLeavey, Christine, and Sutskever, Ilya. "Robust speech recognition via large-scale weak supervision". In: *International conference on machine learning*. PMLR. 2023, pp. 28492–28518.

[30] Salt, Barry. "Reaction time: how to edit movies". In: *New Review of Film and Television Studies 9*, no. 3 (2011), pp. 341–357. DOI: 10.1080/17400309.2011.585865.

[31] Salt, Barry. "Statistical style analysis of motion pictures". In: *Film quarterly 28*, no. 1 (1974), pp. 13–22.

[32] Schatz, Thomas. "Desilu, I Love Lucy and the Rise of Network Television". In: *Making Television: Authorship and the Production Process*, ed. by Robert J. Thompson and Gary Burns. New York: Praeger, 1990.

[33] Soucek, Tomás and Lokoc, Jakub. "Transnet v2: An effective deep network architecture for fast shot transition detection". In: *Proceedings of the 32nd ACM International Conference on Multimedia*. 2024, pp. 11218–11221.

[34] Spigel, Lynn. *Make Room for TV: Television and the Family Ideal in Postwar America*. Chicago: University of Chicago Press, 1992.

[35] Tian, Ying, Liu, Siyun, and Wang, Jianying. "A Corpus Study on the Difference of Turn-Taking in Online Audio, Online Video, and Face-to-Face Conversation". In: *Language and Speech 67*, no. 3 (2024), pp. 593–616.

[36] Tsivian, Yuri. "Cutting and Framing in Bauer's and Kuleshov's Films". In: *Kintop: Jahrbuch zur Erforschung des Frühen Films 1* (1992), pp. 103–113.

[37] Tsivian, Yuri and Civjans, Gunars. "Cinemetrics: Movie Measurement and Study Tool Database". http://www.cinemetrics.lv. Accessed: 2025-07-10. 2005.

[38] Tueth, Michael V. "Fun City: TV's Urban Situation Comedies of the 1990s". In: *Journal of Popular Film and Television 28*, no. 3 (2000), pp. 98–107.

[39] Warner, Kristen J. “In the time of plastic representation”. In: *Film Quarterly 71*, no. 2 (2017), pp. 32–37.

[40] Warner, Kristen J. *The cultural politics of colorblind TV casting*. Routledge, 2015.

[41] Williams, Raymond. *Television: Technology and Cultural Form*. Routledge, 1974.

A Supplemental Data

Series	MSL	1 Speaker F1	1 Speaker F2	1 Speaker F3	2 Speakers F1	2 Speakers F2	2 Speakers F3	3+ Speakers F1	3+ Speakers F2	3+ Speakers F3
Brooklyn 99	2.0	1.2	1.1	0.9	1.4	1.5	1.5	1.5	1.9	2.0
Kim's Convenience	2.1	1.1	1.2	1.1	1.3	1.9	1.6	1.6	2.5	2.6
30 Rock	2.1	1.3	1.2	1.1	1.6	2.0	1.5	1.8	2.3	2.2
Fresh Off The Boat	2.1	1.4	1.2	1.1	1.6	2.0	1.7	1.8	2.5	1.9
Community	2.2	1.4	1.2	1.0	1.7	2.1	1.8	2.0	3.0	2.7
Black-ish	2.2	1.2	1.1	0.9	1.5	1.7	1.6	1.7	2.6	1.9
Parks and Recreation	2.3	1.3	1.1	0.9	1.7	1.9	1.7	2.3	2.6	2.5
Arrested Development	2.3	1.1	1.2	1.2	1.7	2.3	2.1	2.3	3.3	2.9
The Good Place	2.3	1.3	1.3	1.2	1.7	1.9	1.8	1.5	2.2	1.7
How I Met Your Mother	2.3	1.7	1.5	1.3	1.9	2.4	2.0	2.8	3.7	2.9
The Big Bang Theory	2.6	1.9	1.8	1.5	1.9	2.5	2.0	2.2	3.0	2.8
The Office (US)	2.7	1.7	1.5	1.3	2.1	2.8	2.2	3.8	4.5	3.6
Seinfeld	2.8	1.8	1.9	1.9	1.9	3.0	2.8	3.2	4.4	4.5
Friends	2.8	2.0	2.0	1.8	2.1	2.9	2.5	3.2	4.1	3.4
Modern Family	2.9	1.7	1.7	1.4	2.1	3.0	2.5	4.0	4.2	4.6
I Dream of Jeannie	3.0	1.8	2.3	3.0	2.9	8.0	6.1	15.2	17.7	13.2
Cheers	3.2	1.7	1.9	1.7	2.7	3.9	3.4	4.6	6.0	5.7
Bewitched	3.3	2.1	2.7	3.1	3.5	6.6	6.0	8.2	11.0	10.0
Frasier	3.3	1.8	2.0	1.8	2.1	3.1	2.7	3.7	4.2	4.2
Everyb. Loves Raymond	3.4	2.0	2.2	1.9	3.1	4.0	2.9	4.6	6.2	4.9
That '70s Show	3.8	2.5	2.5	2.5	3.0	4.2	3.6	7.0	5.8	5.8
Mary Tyler Moore Show	3.8	2.2	2.7	2.4	2.9	4.4	3.3	5.1	6.4	6.1
Living Single	3.8	2.6	3.0	2.5	3.2	4.8	3.7	5.4	6.3	6.3
My Living Doll	3.9	2.2	2.8	2.2	4.3	8.9	4.5	7.9	16.8	7.9
Donna Reed Show	3.9	2.1	2.7	3.1	3.2	8.8	7.6	15.6	18.1	14.4
Fresh Prince	4.0	2.3	2.9	2.6	3.3	5.2	4.8	7.1	6.9	6.0
Dick Van Dyke Show	4.1	2.0	2.1	2.9	3.5	4.7	4.1	8.9	7.8	7.0
I Love Lucy	4.4	2.1	2.7	3.3	2.7	4.2	4.0	4.6	6.9	7.4
Good Times	4.5	2.8	3.0	2.7	3.2	5.2	4.5	5.6	6.9	7.4
Sanford and Son	5.1	2.9	3.7	3.3	3.6	6.2	4.4	7.3	8.8	7.3
All in the Family	5.3	3.1	3.4	3.3	4.0	5.8	5.0	8.1	8.7	7.7

Table 6: The normal-corrected median absolute deviation (MAD) values corresponding to the median values in Table 3. The overall MSL (in seconds) is included in the first column because it was used to sort the table, which is aligned with the results in Tables 3–4.