Soci@lCommunication

Social Communication

Soci@lCommunication ONLINE JOURNAL

ISSN: 2450-7563 (Print) (Online) Journal homepage: journals.wsiz.edu.pl/sc

Quantifying Modal Density in Climate Change Discourse: A Multimodal Analysis

Gaëlle Ferré

To cite this article: Ferré, G. (2025). Quantifying Modal Density in Climate Change Discourse: A Multimodal Analysis. *Social Communication. Online Journal, 2*(27), 91-102.
To link to this article: https://doi.org/10.57656/sc-2025-0019

Published online: 31 Dec 2025

UNIVERSITY of INFORMATION TECHNOLOGY and MANAGEMENT
in Rzeszow, POLAND

SOCIAL COMMUNICATION
2025, Vol.27 NO. 2, ISSN 2450-7563
https://doi.org/10.57656/sc-2025-0019

Soci@lCommunication ONLINE JOURNAL

Quantifying Modal Density in Climate Change Discourse: A Multimodal Analysis

Gaëlle Ferré University of Poitiers, France, gaelle.ferre@univ-poitiers.fr

Multimodal Discourse Analysis, prosody, gesture, multimodal constructs, modal density

ABSTRACT

This paper examines modal density, a concept introduced in Discourse Analysis by Norris [2004, 2009], in news media. Modal density combines intensity (modal prominence) and complexity (modal interactions), and is believed to vary across media genres and speech acts in climate change communication. Using data from UCLA's NewsScape corpus, 500 short video clips from genres such as news reports, talk shows, weather forecasts and political speeches were analyzed. The study measured modal intensity and complexity across speech, prosody, and visual resources. Findings highlight differences across genres and speech acts, offering insights into multimodal strategies that promote public engagement with climate change. The study also introduces a quantitative method to compare modal density across genres, which improves our understanding of ecological discourse.

Introduction

Climate change discourse is complex and multidimensional, as it engages scientific, political, social, and ethical perspectives [Fløttum, 2014]. It shapes public opinion, influences policy, mobilizes action, and requires communication to connect with audiences emotionally, scientifically, and ethically [Taylor, 2013; Kumpu, 2022]. While grounded in a scientific narrative on causes and impacts [Dryzek, 2006], the strategies used to convey this narrative vary across media and genres that "define communicative purposes" [Bateman et al., 2017: 131]. In Bateman's framework, genres are not merely textual categories but socially recognised communicative configurations: they represent conventionalised, structured solutions developed by a community to accomplish a specific communicative task. As such, each genre orchestrates rhetorical strategies across the semiotic resources afforded by its medium, simultaneously pursuing its communicative goals and signalling its own identity to its audience [Bateman, 2008][1]. To understand the effectiveness of climate change discourse, attention must be paid to how it is delivered through speech, visuals, and prosody, which shape its impact. Research in multimodal discourse analysis (MDA) has explored how different modes contribute to communication [Norris, 2004, 2009], but the concept of modal density—combining modal intensity and complexity—remains underutilized in this scientific approach.

This paper addresses the research question: how can we quantify differences in modal density in climate change discourse? Modal density, as defined by Norris [2004, 2009], increases when multiple modes are employed in intricate ways, which enhances both intensity and complexity. For example, political speeches may increase density through high-intensity prosody and visuals, while scientific reports rely on complex data visualizations. By quantifying modal intensity and complexity, this study offers insights into how different genres of climate change communication, such as news reports and political speeches, balance these elements. This approach also aligns with Siefkes' [2015] concept of *intersemiosis*, which involves *intermodal relation types* and "context-sensitive influences between different strata of semiotic modes". Intersemiosis captures "the processes through which semiotic modes influence one another and shape the formal, semantic, and/or stylistic structure of a multimodal text in a definable way" [2015: 117], underscoring the importance of examining these interactions rather than treating modes independently.

The study uses a mixed-methods approach, drawing on the NewsScape corpus from UCLA's NewsScape Library, to analyze shifts in modal intensity and complexity across genres. This analysis helps explore the relationship between modal density and communicative

[1]See also the Glossary of multimodal terms: Genre. (n.d.). Retrieved from https://multimodalityglossary.wordpress.com/genre/ (last accessed November 12, 2025).

effectiveness. Modal configurations that balance intensity and complexity may better engage and inform audiences. Political speeches with emotionally charged language and visuals may drive urgency, while scientific discourse may focus on complexity. This research is part of a broader project on climate change discourse, including studies on discourse styles [Ferré, 2024] and metaphors in visual communication, with a particular focus on climate change and COVID-19 [Ferré & Attou, 2025], a study that was very much in line with Dancygier [2023]. Together, these studies aim to provide a comprehensive understanding of how linguistic and visual strategies shape public perception and action on climate change.

Modal density

Norris' [2004, 2009] concept of modal density refers to the intensity and presence of different modes (e.g., visual, textual, gestural) in a communicative context. It shows how modes contribute to overall meaning. This contrasts with Kress and van Leeuwen's visual grammar [1990], which focuses on semiotic choices like colour and composition in visual modes. While Kress and van Leeuwen examine individual modes, Norris emphasizes their interaction and cumulative impact on multimodal texts. Other approaches, such as Bateman's multimodal interactional analysis [2008], focus more on syntactic relationships between modes, often with less attention to overall modal density.

Norris defines modal density as a combination of modal intensity (the strength of a mode) and modal complexity (the intricacy of interaction between modes). Although early research was primarily qualitative, recent advances call for more systematic, quantitative measures of these features across multimodal corpora and genres.

Quantitative methods are increasingly used to analyse multimodal elements objectively. For instance, Pagel et al. [2024] developed tools to assess prosodic prominence in speech using acoustic analysis with software like Praat [Boersma and Weenink, 2022], while Bezemer and Kress [2009] explored visual complexity in educational media. However, few studies have combined these metrics into a unified measure of modal density across modes.

This study aims to fill that gap by offering a quantitative method for the measurement of modal density. Adopting the perspective of multimodal analysis, it quantifies modal intensity and complexity using a scoring system for speech, prosody, and visual elements. For intensity, we include measures like prosodic emphasis and frequency of high-degree words, while complexity is assessed through technical language, visual layering, and gestural intricacy. We calculate modal density scores for each mode and compare them across genres and speech acts.

Building on Norris' framework, this research integrates quantitative methods to compare modal density across larger media corpora. This enhances the objectivity of multimodal analysis and offers insights into how multimodal features are strategically used in climate change communication. The study elaborates on modal density features in speech, prosody, and visuals, drawing from prior work on salience [Landragin, 2011; Boswijk and Coler, 2020; Ferré, 2014, 2024], which directly influences modal density.

Modal density in speech

Speech can vary in intensity in several ways. Semantically, using gradable adjectives and adverbs that express a high degree (e.g., *very, extremely*) or quantifiers (e.g., *more, a lot of, many*) adds intensity. Certain modal auxiliaries, such as *can, must, will*, and the emphatic auxiliary do, convey higher certainty and increase modal intensity compared to modals like *may, might, or should*. For example, must is often used in persuasive writing to convince the audience [Hansen, 2018]. Additionally, including precise dates and figures strengthens statements and boosts speech intensity.

Example 1 below, with no quantifiers, degree adjectives/adverbs, modal auxiliaries, or precise figures, is far less intense than Example 2. In Example 2, modal auxiliary *will* indicates a high degree of certainty, while *more*, *global*, and the figure *55* add precision and amplify intensity.

(1) This guy may be concerned about global warming because he is drooping[2].

(2016-12-12_1900US_FOX-News_Americas_News_HQ)[3]

(2) The agreement will take effect when it's ratified by 55 dignitaries that account for 55 % or more of global greenhouse gas emissions.

(2016-09-22_1300_US_KCET_Newsroom_Tokyo)

Speech may also vary in terms of complexity. A

higher complexity is achieved with complex syntactic structure (sentences that combine at least two clauses), but also with the use of technical words and acronyms, which are less transparent than other speech items and often require a definition of some sort. The presence of reported speech also introduces more complexity since it highlights several enunciative sources and increases the linguistic polyphony of the message [Fløttum, 2010]. Finally, metaphors have been described as mappings between two domains: a source and a target domain [Lakoff and Johnson, 2003], and can be considered more complex from a cognitive point of view. This is the case in Example 3 for instance, in which the ARGUMENT IS WAR metaphor is referred to with the mention of the word ***battle***.

(3) But in the northwest, there's an important battle over a carbon tax ballot initiative in Washington State next month.

(2016-10-21_0100_US_KOCE_The_PBS_Newshour)

Modal density in prosody

Higher intensity may also be achieved in speech under the form of specific prosodic constructs. The largest part of our linguistic messages is uttered in what is understood as broad focus. The whole utterance is then considered as relevant in the activation state of listeners. Broad focus is generally marked in statements by a regularly decreasing pitch and intensity, a final falling tone, and the last syllable of the Intonation Phrase (IP) is usually longer too [Féry, 2001; Wells, 2006]. In contrast, only part of the linguistic message is considered as relevant in the activation state of listeners or as more important speech content in narrow focus. There is some emphasis on this particular part of speech, which is characterized by higher intensity and pitch [Brenier *et al.*, 2005; Herment-Dujardin and Hirst, 2002; Pagel *et al.*, 2024]. The syllables under emphatic stress are generally lengthened especially their onsets [Astésano *et al.*, 2004]. The more emphatic stresses an utterance contains, the more intense the message. Example 4 is uttered with two emphatic stresses on *change* and on *platform*, as shown in Figure 1 that presents the pitch contour of the utterance designed in the speech analysis tool Praat. The compound *climate change* is preceded by a silent pause which contributes to the perceived emphasis [Strangert, 2003] but only *change* is uttered with a large falling-rising contour. At the end of the utterance, the large falling contour on the word *platform*, which begins at a higher pitch than the rest of the utterance, also contributes to the perception of emphasis.

(4) Ivanka Trump wants to make climate change part of her platform.

(2016-12-22_0837_US_KNBC_Late_Night_With_Seth_Meyers_442-453)

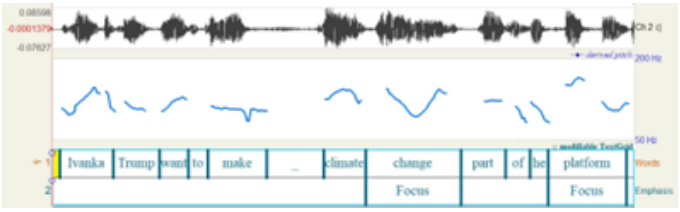


Figure 1 – Waveform, pitch curve and transcription of Example 4 with Praat (Boersma and Weenink, 2022)

Utterances are frequently uttered with more than one emphatic stress (especially in complex sentences) using a speech style which is quite common in news reports [Rodero, 2013]. Some utterances show an even higher degree of emphasis as speakers adopt a particular speech rhythm called *beat prosody*. Simon and Grobet [2005, p. 16] define beat prosody as showing the following acoustic features:

- Recurring prominent syllables perceived as isochronous (identical intervals);
- A higher ratio of stressed to unstressed syllables;
- Sometimes slower speech rate, though not always.

The utterance in Example 5 illustrates what is meant by beat prosody. Slanted lines in the transcription of the example indicate IP boundaries and pitch accents are marked with capital letters. At the beginning of the sentence, IPs align with syntactic grouping so that « it will take vast investment » and « to avoid tragedies like this » have a respective duration of 1.757 and 1.597 sec. The speaker then adopts much shorter IPs, ranging in duration from 0.356 to 0.823 sec, and chunks his speech

[1]The speaker is referring to a snowman.

[2]Each example is systematically followed by its reference in the corpus.

into two to four-syllable units. Four out of the five IPs involved are even followed by silent pauses which also play a role in the perception of emphasis. This part of the sentence—transcribed in square brackets—gives the impression that each lexical item is emphasized with significant stress. Following this, the speaker returns to his original rhythm on « it is still raining », which lasts 1.167 sec.

(5) / Experts WARN / it will take VAST investment / to avoid tragedies like THIS / [if the PAttern / of exTREME / WEAther / conTInues, / and toDAY] / it is still RAIning. /

(2016-08-04_0100_UK_KCET_BBC_World_News_1157-1174)

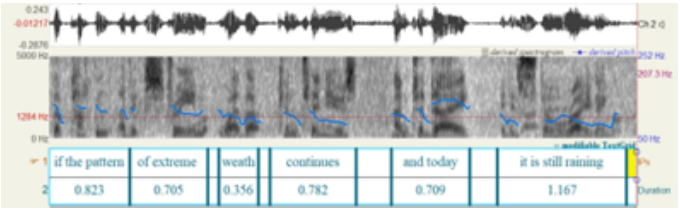


Figure 2 – Rhythmic pattern illustrating beat prosody in Example 5 analysed in Praat (Boersma and Weenink 2022, top: waveform, middle: spectrogram and pitch curve, bottom: transcription and duration of IPs)

Turning to prosodic complexity, the utterance quoted above in Example 4 also shows complexity in pitch contours. The default pattern for statements uttered in a single IP is to have a final falling contour. However, a sentence may be uttered with more than one IP. In this case, a series of falling contours indicate that the IPs are to be considered as independent from one another, whereas a rising contour at the end of one IP of the utterance indicates dependency [Wells, 2006]. This was the case in Example 4 which can therefore be considered as being complex in terms of prosody as well as intense.

Modal density in gestures and visuals

Gestures and visuals finally also contribute to modal density and can add weight to intensity or complexity. Greater intensity is conveyed through the use of three types of bodily movements: head beats and eyebrow raises have been shown to highlight speech items and to make them perceived as more prominent in the speech flow than the rest of the utterance [Al Moubayed *et al*., 2010; Dohen and Lœvenbruck, 2009; Krahmer *et al*., 2002; Krahmer and Swerts, 2007; Swerts and Krahmer, 2008, 2010]. Hand beats play a similar role [Prieto Vives *et al*. 2018; Swerts and Krahmer 2007], while pointing gestures, as demonstrated by Edeline and Klinkenberg [2021], are also associated with the expression of emphasis. However, whereas beats emphasize particular elements of speech, pointing gestures highlight specific locations or concrete referents within that space. These two non-representational gesture types, therefore, add intensity to the speech content.

In Example 6 below, the speaker performs two beat gestures with his right hand, as illustrated in Figure 3, which shows the beginning and end of the first beat stroke, *e.g.* the relevant part of a hand gesture (Kendon, 2004). The second gesture mirrors the first, and together, they highlight the contrast between *the carbon tax and the estate tax* in the speaker's sentence. However, the two gestures do not add any semantic content.

(a)Beginning of beat gesture stroke (b)End of beat gesture stroke

Figure 3 – Non-representational gesture in a political debate (Example 6)

(6) Somebody in the far right could say, gee, I'd love [a carbon tax]^{BEAT1} as long as we use it [to reduce the estate tax]^{BEAT2}.

(2016-10-21_0100_US_KOCE_The_PBS_Newshour_3093-3099)

Representational gestures, which represent some aspects of the objects or actions depicted, add complexity to the message. The affordances of the gestural semiotic mode convey information absent in speech, ensuring that gesture and speech are never entirely redundant, even when they express the duplicate linguistic content.

In Example 7 below, the speaker performs two representational gestures. He first draws his index and thumb fingers together to depict a very small quantity, and then opens up both his arms to depict an extremely large quantity. The two gestures present a stark contrast in size and are not redundant with speech. They add

modal complexity to the speaker's discourse and serve his denial of the effects of climate change.

(7) The hysteria over global warming—it should be [this..., you know, a little bit like this]G1. [It's like this] G2.
(2016-12-21_2200_US_FOX-News_The_Five_1775.03-1786)

(a)Representational gesture 1

(b)Representational gesture 2

Figure 4 – Representational gestures in a political debate (Example 7)

In television programs, other visual resources may also be used to add complexity to speech: text boxes, graphs, still images, and video clips. Figure 5 presents quite a good example of visual complexity. In the top-left corner of the screen is a still image of Barack Obama standing before a group of people. The text in white on the picture summarises the topic, which is further elaborated in the blue box at the bottom of the screen, while the red text box announces forthcoming topics in the news report. The middle section of the screen is primarily occupied by a video clip of drilling platforms at sea, in line with the current topic, while the host is visible on the right side of the screen. This Figure, in which the screen is visually saturated with information, presents a stark contrast with Figure 6, in which only the

Figure 5 – High visual complexity in a news report (2016-1221_2100_US_MSNBC_ MSNBC_Live_ With_Steve_Kornacki)

Figure 6 – Low visual complexity in a debate on climate change (2016-0908 _0100_US_KOCE_The_ PBS_Newshour)

host is visible and this can be considered as presenting a very low degree of visual complexity.

Data and methodology

Corpus

This study uses data from UCLA's NewsScape corpus, part of the Distributed Little Red Hen Lab™, co-directed by Francis Steen and Mark Turner. The corpus contains over 200 million words from transcribed television broadcasts like news, talk shows, political speeches, and weather forecasts. For this analysis, 500 video clips each showing a sentence relevant to climate change discourse were selected using keywords such as *climate change, global warming*, and *extreme weather*. Using CQPweb [Hardie, 2012] (Figure 7), the clips were cleaned of duplicates and irrelevant content, then exported for annotation with Uhrig's [2018] Rapid Annotator (Figure 8). Research suggests that brief, focused video segments align with the average attention span and cognitive processing abilities, which enables viewers to process and integrate information more effectively [Guo *et al*., 2014; Mayer, 2009]. The choice of short clips was therefore motivated by the need to align with this attention span while still capturing salient moments for multimodal analysis.

Annotation p ocess

The 500 clips were annotated across three semiotic modes: speech, prosody, and visuals. Each clip was coded for various features that contribute to modal intensity and modal complexity. The annotations followed a

[4]With the exception of the channel logo which appears in nearly all the videos in the corpus and was excluded from the study.
[5]<http://redhenlab.org> [last accessed January 2, 2025].
[6]These were the only genres available in the corpus for the search hits used as material in the present study, although other genres would have been of interest as well. In a previous study [Ferré, 2023], various online videos such as TED Talks and educational materials were used; however, such sources are difficult to integrate into a quantitative approach because they do not yield a sufficiently high number of comparable occurrences for patterns to emerge and because they cannot be annotated with the tools used here.

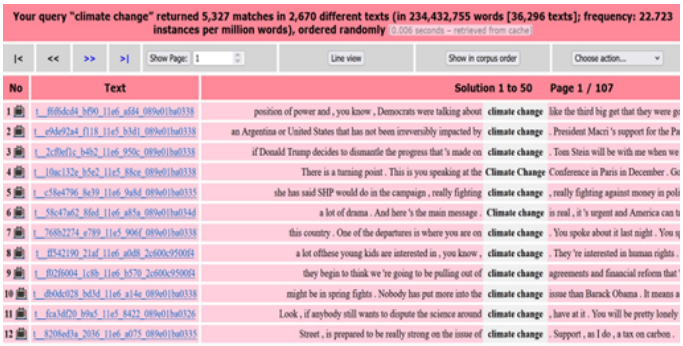


Figure 7 – The NewsScape query concordancer shown in CQPweb (Hardie, 2012) for the corpus collected by the Distributed Little Red Hen Lab

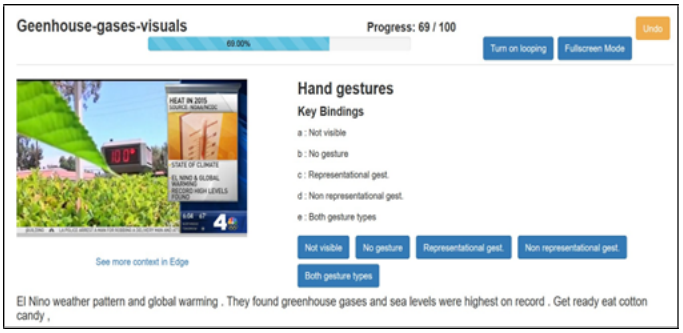


Figure 8 – The rapid annotator developed by Uhrig (2018) for the Distributed Little Red Hen Lab

predefined coding scheme that assigned specific values to each feature of interest. These values are detailed in Table 1 in the appendix.

Quanti ying Modal Intensity

Modal intensity, which once again refers to the prominence or strength of a particular mode, is quantified in this study using the following criteria, defined in a previous qualitative study mentioned before [Ferré, 2024]:

Speech: Intensity in speech was measured through the use of high-degree words (e.g., *more, very, extremely*), modal auxiliaries (e.g., *must, will*), and specific figures or dates that provide precision and assertiveness. For each occurrence of these elements, a score of 1 was assigned. For instance, the sentence «55 countries will ratify the agreement» would receive a score of 1 for the modal auxiliary *will* and 1 for the specific figure *55*.

Prosody: Prosodic intensity was measured based on the perception of emphatic stress, with increasing levels assigned according to the number and prominence of emphatic stresses. Sentences with no emphatic stress were scored 0, sentences with one emphatic stress scored 1, two emphatic stresses scored 2, and so on. Additionally, beat prosody (a series of emphatic stresses in speech) was given the highest score of 3.

Visuals: Visual intensity was quantified by assessing the presence of head beats, eyebrow raises, and non-representational gestures (e.g., points and beats), each contributing to the prominence of the visual mode. Each gesture or bodily emphasis counted as 1 point, and a higher cumulative score indicated greater visual intensity.

Quanti ying Modal Complexity

As already said above, modal complexity refers to the intricacy of the interaction between modes and therefore contributes to the cognitive effort required to process the discourse. It was measured using the following features:

Speech: Complexity was determined by syntactic structure; simple sentences (containing one clause) were scored 0, and complex sentences (containing multiple clauses or embedded structures) were scored 1. The presence of technical terms, acronyms, or metaphors also contributed to higher complexity scores, with each technical or metaphorical element contributing 1 point.

Prosody: Prosodic complexity was measured based on pitch contours and the use of multiple IPs. Simple falling or rising contours were scored 0, while more complex combinations of rising and falling contours across multiple IPs were scored 1 or 2, depending on the level of intricacy.

Visuals: Visual complexity was assessed based on the use of representational gestures (e.g., gestures that visually depict concrete or abstract concepts like *global warming*), text overlays (e.g., graphs, diagrams, or captions that accompany the visual content), and the presence of multiple visual layers (e.g., video clips, still images, and text simultaneously). Each representational gesture, text box, or additional visual layer added 1 point to the visual complexity score.

Aggregation of Modal Density S ores

After individual scores for intensity and complexity in each mode were assigned, they were aggregated to produce a total modal density score for each short video clip. This was done separately for intensity and complexity in each mode, and then combined to produce an overall modal density score. For example, if a video clip had the following scores:

— Speech intensity: 3 — Speech complexity: 2
— Prosodic intensity: 2 — Prosodic complexity: 1
— Visual intensity: 1 — Visual complexity: 3

The total intensity score would be 3 + 2 + 1 = 6, and the total complexity score would be 2 + 1 + 3 = 6. The final modal density score for the clip would then be the sum of both, resulting in 12.

In MDA, no single mode is inherently more important or takes precedence over another. Each mode—whether visual, textual, or prosodic—contributes uniquely to the overall communicative meaning. In this study, the scoring system for modal intensity and complexity does not assign greater weight to any specific mode; rather, it evaluates how modes interact and contribute to the message's overall impact. This approach reflects the MDA principle, in which meaning emerges from the interplay between modes rather than from any single mode in isolation. Thus, the scoring system ensures that all modes are considered in their respective roles without prioritizing one over the other, which is in line with Adami [2017].

Results

The Kruskal-Wallis tests conducted on the data showed significant differences in overall modal density across both genres ($p < 0.01$) and speech acts ($p < 0.01$). This confirms that the interaction between intensity and complexity is genre-specific and reflects the communicative goals of different discourse types. Pairwise comparisons further demonstrated that political speeches had significantly higher intensity scores compared to news reports ($p = 0.002$), while narratives consistently outscored fact statements in both complexity and intensity ($p = 0.001$). This is shown in the boxplots and graphs in Figures 9-13, while scores are given in Tables 2 & 3.

Modal Density Across Genres

News reports exhibited high modal complexity but lower modal intensity, with a total modal density score of 7.0. This is attributed to the detailed explanations and the frequent use of technical terms and visual aids such as graphs and charts. While the content is complex, the intensity—both prosodic and visual—is moderate, as news broadcasts often aim for neutrality and avoid emotional appeals.

Political Speeches: Political speeches demonstrated similar levels of modal intensity and complexity, with a total modal density score of 7.0. The frequent use of emphatic stress caused the elevated intensity, high-degree modal auxiliaries (e.g., *must, will*), and gestures such as hand beats and head movements. However, political speeches also maintained moderate complexity, balancing the need for urgency and emotion with sufficient detail.

Talk Shows: With a total modal density score of 7.2, talk shows exhibited moderate visual intensity, as indicated by a 1.7 score for visuals. This reflects the dynamic use of gestures and facial expressions typical of the genre. However, the overall complexity was moderate, with a modal complexity score of 3.6. While talk shows aim to engage and entertain the audience, the complexity remains relatively low, as the discussions often focus on conversational, rather than deeply analytical, content. This balance between interaction and entertainment results in moderate complexity across speech, prosody, and visuals.

Weather forecasts displayed moderate modal complexity (score of 3.7) and relatively low modal intensity (score of 2.7), with a total modal density of 6.4. This reflects their data-focused, straightforward nature, which often relies on visual elements such as maps and charts to convey weather patterns. Since they deliver factual information, the intensity—both prosodic and visual—remains relatively low, as indicated by the 1.3 score for prosody and 1.0 for visuals. This suggests that while extreme weather events may be covered, the goal of weather forecasts is to maintain neutrality. They thus avoid dramatic or overly persuasive tones, even when they address significant events such as storms or heatwaves, often linked to climate change.

In sum, these results illustrate how modal density shifts across genres. Weather forecasts emphasize moderate complexity (3.7) through the use of visual aids like maps and charts, but maintain lower intensity (2.7), as they focus on delivering clear and factual information. News reports, on the other hand, prioritize high complexity (4.1) but with moderate intensity (2.9). These programs offer detailed explanations and maintain a neutral tone. Political speeches balance both intensity (3.4) and complexity (3.6), evoking urgency and emotional responses, while still delivering a moderate level of detail. Talk shows, similarly, strike a balance between moderate complexity (3.6) and visual intensity (3.6), using dynamic visuals, gestures, and conversational styles to captivate and engage their audience.

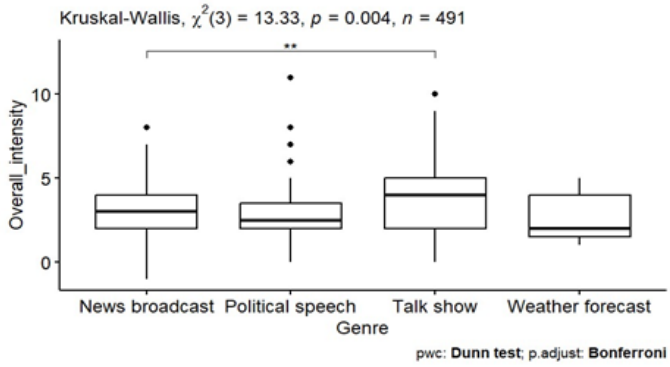


Figure 9 – Differences in overall intensity between genres in the NewsScape corpus

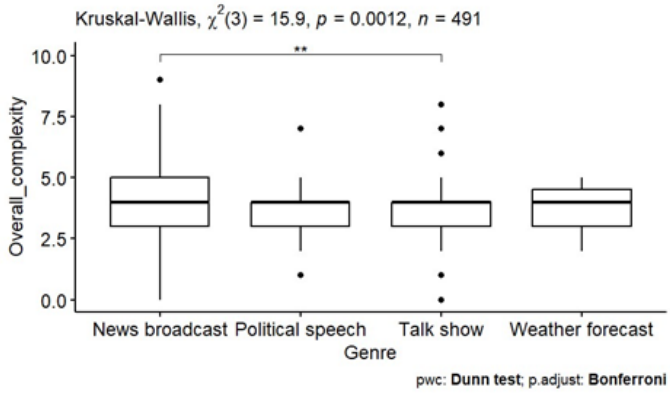


Figure 10 – Differences in overall complexity between genres in the NewsScape corpus

Modal Density Across Speech Acts

The study also found notable differences in modal density across various speech acts as shown in Table 3 in the Appendix, which illustrates how communicative objectives influence the interaction between intensity and complexity. While not all speech acts will be described in detail, the focus will be on the most significant ones, which highlight major patterns in how modal density operates across different communicative contexts.

Contradiction has one of the highest modal density scores (7.1), which is composed of a modal complexity score of 4.2 and a modal intensity score of 2.9. This high density reflects the intricate interaction of modes used when contradictions are expressed, as they often involve complex language, multiple enunciative sources, nuanced prosodic patterns, and visual elements such as facial expressions or gestures to emphasize opposing viewpoints. The greater complexity stems from the need to articulate and balance contrasting ideas, while the moderate intensity underscores the importance of highlighting these opposing perspectives without overwhelming the listener.

Narrative speech acts demonstrated a high overall modal density, with a total score of 7.3, largely due to their use of complex sentence structures, technical language, and representational gestures that enrich the narrative. These elements, combined with the use of visual aids such as video clips, increased both intensity and complexity. Narration can be considered a highly informative speech act. The complexity score for narration was 4.4, with notable contributions from speech (1.6), prosody (1.5), and visuals (1.3). The intensity score was 2.9, with speech and prosody contributing 0.9 and 1.2, respectively, while visuals played a smaller role.

Incitative speech acts, such as calls to action in political speeches, had a total modal density of 6.4. These acts were marked by moderate intensity, driven by emotionally charged language, emphatic stress, and gestures intended to motivate the audience. While complexity was slightly lower, reflecting a focus on urging immediate action rather than providing detailed explanations, the overall balance of intensity and complexity made the incitations effective at prompting engagement. It is surprising that these speech acts are not more intense, given their potential to drive action and influence audiences, yet they still maintain a significant impact with their measured use of intensity and complexity.

Descriptive speech acts had a mean modal density of 7.0. These acts were characterized by moderate complexity, driven by detailed language, clear structure, and often accompanied by visuals to enhance the imagery. While the intensity was lower than that of other speech acts, rich descriptions effectively convey information and create vivid mental images for the audience.

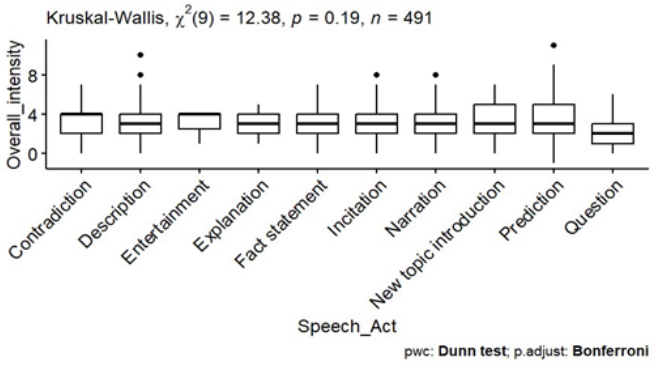


Figure 11 – Differences in overall intensity between speech acts in the NewsScape corpus

The differences in modal density across speech acts highlight how intensity (Figure 11) and complexity (Figure 12) are strategically deployed to achieve various communicative goals.

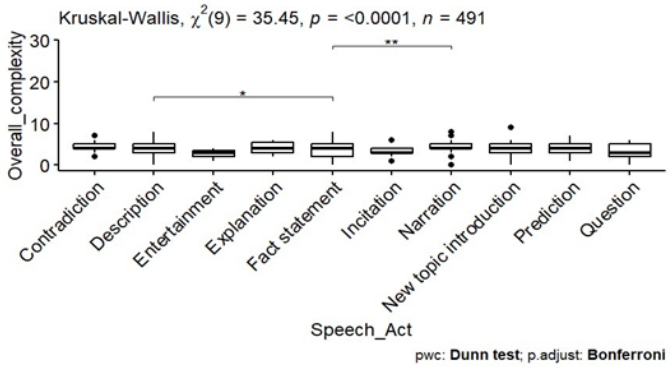
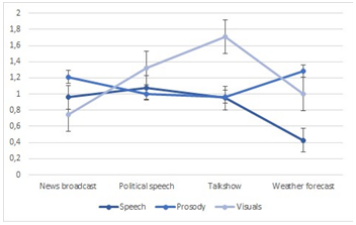
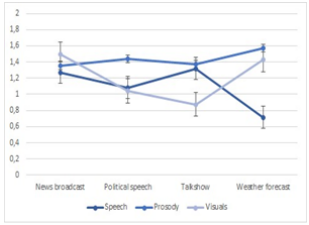


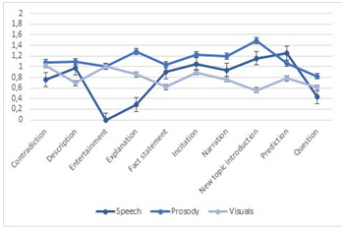
Figure 12 – Differences in overall complexity between speech acts in the NewsScape corpus



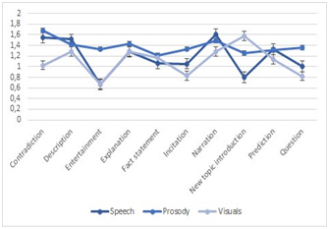
(a) Mean genre intensity scores



(b) Mean genre complexity scores



(c) Mean speech act intensity scores



(d)Mean speech act complexity scores

Figure 13 – Mean intensity and complexity scores across genres and speech acts

Discussion

The results of this study highlight modal density as a key factor in shaping climate change discourse across different genres and speech acts. This research thus offers objective insights into how multimodal elements interact to create discourse that is both informative and persuasive. This approach is particularly relevant in climate change communication, where effective discourse can influence public understanding, engagement, and action.

Genre-Based Discourse Strategies in Climate Change Communicatio

In climate change discourse, it is crucial to balance modal intensity (emotional strength) and modal complexity (informational richness). High-intensity messages create urgency, while high-complexity ones add credibility. This balance is vital to counter misinformation, emotional fatigue, and the complexity of climate science.

Political speeches and activist messages use a strategic blend of intensity and complexity to emphasize urgency while maintaining credibility. Emotional language, stress, and gestures aim to evoke responsibility or hope, but excessive emotion tends to reduce credibility. Balancing these dimensions is important to both incite audiences to act without alienating them.

Weather forecasts, with lower intensity, present extreme weather visually, associating it with climate change. They focus on clear, factual visuals like maps which ensures they remain informational rather than persuasive. While their emotional impact is subtle, the dramatic visuals can still raise awareness of climate risks.

News reports prioritize high complexity with lower intensity, as they use technical language and visuals like graphs to inform and educate. Their focus on neutrality limits emotional appeal but is crucial against misinformation. To boost engagement, these reports may need to be supplemented by more emotionally engaging content.

Talk shows offer moderate intensity and complexity, making climate change accessible through conversational formats. They blend relatable dialogue with visuals, and thus bridge the gap between purely informative and emotionally charged discourse.

Modal density needs to be adapted to context and audience in order to overcome feelings of disengagement and skepticism. The study suggests that different genres, from high-intensity political speeches to fact-driven news reports, play complementary roles in mobilizing effective climate responses.

Implications or Various Speech Acts

A central challenge in climate change communication consists in overcoming public disengagement. Overly complex messages may alienate audiences, while overly emotional ones may provoke skepticism, which is why a good balance is preferable.

Narratives are particularly effective as they balance intensity and complexity, and score high in modal density. Documentaries and public awareness campaigns use relatable narratives grounded in scientific facts to draw people into the discussion.

Contradictions in climate change discourse, which emphasize opposing viewpoints, often reflect moderate modal density, marked by complex interactions of language, multiple enunciative sources, and multimodal elements.

Incitative messages, conveyed by political speeches or social media advocacy, rely on moderate intensity to motivate action. However, oversimplification can alienate informed audiences, so balancing emotional appeal with scientific detail is important.

Descriptive messages show moderate intensity to present detailed information in an accessible way. By balancing clear language and relevant visuals, these messages help convey complex concepts without overwhelming the audience.

Conclusion

This study has demonstrated the value of quantifying modal density, defined by Norris [2004, 2009] as the interaction of modal intensity (strength of a mode) and modal complexity (intricacy of multimodal interaction). This is especially relevant in climate change discourse. The study presented a quantitative framework that sheds light on how multimodal constructs function across genres and speech acts in climate communication. Results showed that varying levels of intensity and complexity are strategically used across media to engage, inform, and persuade. This approach provides a more objective understanding of how climate change discourse shapes public perception and action.

The study applied quantitative measures in MDA, which goes beyond subjective discourse interpretations. Scores for intensity and complexity in each semiotic mode (speech, prosody, visuals) were aggregated into a modal density score, enabling systematic comparisons of discourse styles. This precise analysis of how modes combine to achieve communicative goals complements previous qualitative work [Ferré, 2024].

Understanding the intricacies of modal density in terms of intensity and complexity can improve communication strategies. The study highlights that balanced high modal density makes climate change discourse more engaging, while imbalances between intensity and complexity can reduce its effectiveness. Too much intensity can oversimplify the message, while too much complexity can make it overly technical and fail to capture attention. Achieving the right balance is therefore essential for effective communication.

Quantifying the contributions of semiotic modes provides a better understanding of multimodal constructs and opens new research avenues in MDA that may apply to high-stakes communication, such as public health or political campaigns. Future refinements, such as automated tools for measuring prosodic and visual features, could enable larger-scale analyses, while expanding datasets to include diverse media genres and additional parameters that might be introduced into the study (such as colour or sentence length). They would also provide insights into how modal density adapts to different platforms and influences audience engagement.

Acknowledgements

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. I would like to express my deepest thanks to Francis Steen and Mark Turner, as well as to Peter Uhrig, for giving me access to the Little Red Hen Lab corpus and tools. I would also like to thank the two anonymous reviewers, as well as several participants at the 12ICOM International Conference in Groningen (October 2025), for their positive feedback and insightful comments on a previous version of this paper.

References

Adami E. (2017) Multimodality. [In:] García O, et al. (eds.) The Oxford Handbook of Language and Society. Oxford: OUP, pp. 451–472.

Al Moubayed S., Beskow J. & Granström B. (2010) Auditory visual prominence. From intelligibility to behavior. Journal of Multimodal User Interfaces 3: 299–309.

Astésano C., Magne C., Morel M., Coquillon A., Espesser R., Besson M. & Lacheret-Dujour A. (2004) Marquage acoustique du focus contrastif non codé syntaxiquement en français. [In:] Actes des Journées d'Étude sur la Parole (JEP). Fès, Maroc, pp. 1–4.

Bateman J.A. (2008) Multimodality and Genre. A Foundation for the Systematic Analysis of Multimodal Documents. New York: Palgrave Macmillan.

Bateman J., Wildfeuer J. & Hiippala T. (2017) Multimodality. Foundations, Research and Analysis. A Problem-Oriented Introduction. Berlin, Boston: De Gruyter Mouton.

Bezemer J. & Kress G. (2009) Visualizing English: A social semiotic history of a school subject. Visual Communication 8: 247–262.

Boersma P. & Weenink D. (2022) Praat: doing phonetics by computer (Version 6.2.22) [Computer program].

Boswijk V. & Coler M. (2020) What is Salience? Open Linguistics 6(1): 713– 722.

Brenier J, Cer D.M. & Jurafsky D. (2005) The detection of emphatic words using acoustic and lexical features. [In:] Proceedings of Interspeech 2005. Lisbon, Portugal, pp. 3297–3300.

Dancygier B. (2023) Multimodal media: Framing climate change. Discourse Studies 25, 220-236.

Dohen M. & Lœvenbruck H. (2009) Interaction of audition and vision for the perception of prosodic contrastive focus. Language and Speech 52(2-3): 177–206.

Dryzek J. (2006) The Politics of the Earth: Environmental Discourses. Oxford: OUP.

Edeline F. & Klinkenberg J.M. (2021) L'index. Un dispositif sémiotique puissant et méconnu. [In:] Bertrand D. & Darrault-Harris I. (eds.) A même le sens. Hommage à Jacques Fontanille. Limoges: Lambert Lucas, pp. 253–263.

Ferré, G. (2014) A Multimodal Approach to Markedness in Spoken French. Speech Communication 57, 268-282.

Ferré, G. (2024) Exposition et exhortation : deux facettes du discours écologique. Anglophonia 37, 1-21.

Ferré, G., Attou, R. (2025) Voir la crise en images : métaphores visuelles et cognition dans les représentations du climat et de la pandémie. Corela 23, 1-26.

Fløttum K. (2010) A linguistic and discursive view on climate change discourse. ASp 58: 19–37.

Fløttum K. (2014) Linguistic mediation of climate change discourse. ASp 65: 7–20.

Féry C. (2001) Focus and Phrasing in French. [In:] Féry C. & Sternefeld W. (eds.) Audiatur Vox Sapientiae. A Festschrift for Arnim von Stechow. Berlin: Akademie-Verlag, pp. 153–181.

Guo P.J., Kim J. & Rubin R. (2014) How video production affects student engagement: an empirical study of MOOC videos. [In:] Proceedings of the First ACM Conference on Learning @ Scale Conference. New York, USA: Association for Computing Machinery, p. 41–50.

Hansen B. (2018) Corpus Linguistics and Sociolinguistics. A Study of Variation and Change in the Modal Systems of World Englishes. Leiden: Brill.

Hardie A. (2012) CQPweb — combining power, flexibility and usability in a corpus analysis tool. International Journal of Corpus Linguistics 17(3): 380–409.

Herment-Dujardin S. & Hirst D. (2002) Emphasis in English: a Perceptual Study based on Modified Synthetic Speech. [In:] Bel B. & Marlien I. (eds.) Proceedings of Speech Prosody 2002. Aix en Provence: LPL, pp. 379–382.

Kendon A. (2004) Gesture. Visible Action as Utterance. Cambridge: CUP.

Krahmer, E., Ruttkay, Z., Swerts, M. & Wesselink, W. (2002) Pitch, Eyebrows and the Perception of Focus. [In:] Bel B. & Marlien I. (eds.) Proceedings of Speech Prosody 2002. Aix en Provence: LPL, pp. 443–446.

Krahmer E. & Swerts M. (2007) The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. Journal of Memory and Language 57: 396–414.

Kress G. & van Leeuwen T. (1990) Reading Images. Geelong, Victoria: Deakin University Press.

Kumpu V. (2022) What is public engagement and how does it help to address climate change? A review of climate communication research. Environmental Communication 16(3): 304–316.

Lakoff G. & Johnson M. (2003) Metaphors we live by. Chicago: The University of Chicago Press.

Landragin F. (2011) De la saillance visuelle à la saillance linguistique. [In:] Inkova O. (ed.) Saillance. Aspects linguistiques et communicatifs de la mise en évidence dans un texte. Besançon: Presses Universitaires de Franche-Comté, pp. 67–84.

Legagneux P., et al. (2018) Our house is burning: Discrepancy in climate change vs. biodiversity coverage in the media as compared to scientific literature. Frontiers in Ecology and Evolution 5: 1–6.

Mayer R.E. (2009) Multimedia learning. Cambridge: CUP.

Norris S. (2004) Analyzing Multimodal Interaction. A Methodological Framework. New York, London: Routledge.

Norris S. (2009) Modal density and modal configurations: Multimodal actions. [In:] Jewitt C. (ed.) Routledge Handbook for Multimodal Discourse Analysis. New York, London: Routledge, pp. 78–90.

Pagel L., Roessig S. & Mücke D. (2024) The encoding of prominence relations in supra-laryngeal articulation across speaking styles. Laboratory Phonology 15(1): 1–55.

Prieto Vives P., Cravotta A., Kushch O., Rohrer P.L. & Vilà-Giménez I. (2018) Deconstructing beat gestures: a labelling proposal. [In:] Proceedings of Speech Prosody 2018. Poznan, Poland, pp. 201–205.

Rodero E. (2013) Peculiar styles when narrating the news: The intonation of radio news bulletins. Estudios sobre el Mensaje Periodístico 19: 519–532.

Siefkes M. (2015) How Semiotic Modes Work Together in Multimodal Texts: Defining and Representing Intermodal Relations. Media Linguistics. 10plus1: Living Linguistics 1, 113-131.

Simon A.C. & Grobet A. (2005) Interprétation des scansions rythmiques en français. [In:] Actes du colloque Interface Discours Prosodie. Aix en Provence, pp. 1–19.

Strangert E. (2003) Emphasis by pausing. [In:] Proceedings of 15th ICPhS Barcelona. Barcelona, pp. 2477–2480.

Swerts M. & Krahmer E. (2007) Acoustic Effects of Visual Beats. [In:] Proceedings of the Auditory Visual Speech Processing Conference (AVSP). Hilvenbareek, The Netherlands, paper L4-2.

Swerts M. & Krahmer E. (2008) Facial expression and prosodic prominence: Effects of modality and facial area. Journal of Phonetics 36: 219–238.

Swerts M. & Krahmer E. (2010) Visual prosody *of newsreaders:*

Effects of information structure, emotional content and intended audience on facial expressions. Journal of Phonetics 38(2): 197–206.

Taylor C. (2013) The Discourses of Climate Change. [In:] Cadman T. (ed.) Climate Change and Global Policy Regimes. London: Palgrave Macmillan UK, pp. 17–31.

Uhrig P. (2018) NewsScape and the Distributed Little Red Hen Lab – A digital infrastructure for the large-scale analysis of TV broadcasts. [In:] Zwierlein A.J., et al. (eds.) Anglistentag 2017 in Regensburg. Proceedings of the Conference of the German Association of University Teachers of English. Trier: Wissenschaftlicher Verlag Trier, pp. 99–114.

Wells C.J. (2006) English Intonation. An Introduction. Cambridge: CUP.

Appendix

Table 1 – Annotation levels and values coded in the study as well as their contribution to modal intensity, modal complexity or both (brackets show the number of occurrences for each value)

Level	Annotation values	Modal density
Text	Text genre: *news broadcast* (316), *talk show* (145), *political speech* (25), *weather forecast* (7), *unidentified* (7)	
	Speech act: *Contradiction* (37), *Description* (121), *Entertainment* (3), *Explanation* (7), *Fact statement* (100), *Incitation* (18), *Narration* (63), *New topic introduction* (70), *Prediction* (51), *Question* (28), *Unidentified* (2)	
	High degree words	Intensity and complexity
	Quote: *Yes* (68), *No* (432)	Complexity
	Technical words: *Yes* (97), *No* (403)	Complexity
	Acronyms: *Yes* (41), *No* (459)	Complexity
	Auxiliary	Intensity
	Date/figure: *Yes* (87), *No* (413)	Intensity
	Syntax: *Simple* (137), *Complex* (363)	Complexity
Prosody	Prosodic emphasis: *No emphasis* (145), *Emphasis+* (145), *Emphasis++* (205), *Beat prosody* (5)	Intensity
	Pitch contours: *Fall* (30), *Rise* (0), *Rise-Fall/Fall-Rise* (210), *Rise-Rise* (2), *Fall-Fall* (258)	Complexity
Gestures and visuals	Hand gestures: *Not visible* (272), *No gesture* (46), *Representational gesture* (3), *Non-representational gesture* (156), *Both gesture types* (25)	Intensity and complexity
	Head beat: *Yes* (245), *No* (85), *Not visible* (170)	Intensity
	Eyebrow raises: *Yes* (110), *No* (214), *Not visible* (176)	Intensity
	text box: *Yes* (318), *No* (182)	Complexity
	Graph: *Yes* (27), *No* (473)	Complexity
	Still image: *Yes* (96), *No* (404)	Complexity
	Video clip: *Yes* (170), *No* (330)	Complexity

Table 2 – Complexity and intensity scores across genres in the NewsScape corpus

	MEAN COMPLEXITY			MODAL	MEAN INTENSITY			MODAL
	Speech	Prosody	Visuals	COMP	Speech	Prosody	Visuals	INT
News broadcast	1,3	1,4	1,5	4,1	1,0	1,2	0,7	2,9
Political speech	1,1	1,4	1,0	3,6	1,1	1,0	1,3	3,4
Talk show	1,3	1,4	0,9	3,6	1,0	1,0	1,7	3,6
Weather forecast	0,7	1,6	1,4	3,7	0,4	1,3	1,0	2,7

Table 3 – Complexity and intensity scores across speech acts in the NewsScape corpus

	MEAN COMPLEXITY			MODAL	MEAN INTENSITY			MODAL	MODAL
	Speech	Prosody	Visuals	COMP	Speech	Prosody	Visuals	INT	DENSITY
Contradiction	1,5	1,7	1,0	4,2	0,8	1,1	1,0	2,9	7,1
Description	1,5	1,4	1,3	4,2	1,0	1,1	0,7	2,8	7,0
Entertainment	0,7	1,3	0,7	2,7	0,0	1,0	1,0	2,0	4,7
Explanation	1,3	1,4	1,3	4,0	0,3	1,3	0,9	2,4	6,4
Fact statement	1,1	1,2	1,2	3,4	0,9	1,0	0,6	2,6	6,0
Incitation	1,1	1,3	0,8	3,2	1,1	1,2	0,9	3,2	6,4
Narration	1,6	1,5	1,3	4,4	0,9	1,2	0,8	2,9	7,3
New topic intro.	0,8	1,3	1,6	3,6	1,2	1,5	0,6	3,2	6,8
Prediction	1,3	1,3	1,1	3,8	1,3	1,1	0,8	3,1	6,9
Question	1,0	1,4	0,8	3,2	0,4	0,8	0,6	1,9	5,0