DE GRUYTER MOUTON

Jinyi Huang and Jinjun Wang*

Evidentiality in science from specialization to popularization: A case study of COVID-19 texts

Abstract: Since the beginning of the COVID-19 pandemic outbreak, medical texts on the pandemic have enjoyed wide popularity, and one of the key issues has always been the accuracy and dependability of the information they contain. The use of evidentiality, a linguistic system which functions to indicate the source and credibility of information, is thus worth exploring in COVID-19 texts. Adopting a synthesized framework within the overall model of systemic functional linguistics, this paper sets out to investigate the lexicogrammar and semantics of evidentiality on the basis of data collected in the form of both specialized and popular texts on COVID-19. Evidentiality in these texts is explored along four dimensions: (i) EVIDENTIAL TAXONOMY, where specialized texts favor reporting, while popular texts favor belief and inferring; (ii) INFORMATION SOURCE, where specialized texts highlight the voices of authorship, original research, and patients, whereas popular texts highlight the voices of scientists, institutions, countries, and laypeople; (iii) MODALIZATION, where specialized texts typically indicate a higher degree of modal responsibility than their popular counterparts; and (iv) ENGAGEMENT, where specialized texts favor dialogic expansion and popular texts favor contraction. It is hoped that these findings will shed light on linguistic variation according to different contextual configurations, as well as clarifying rhetorical conventions in discourse communities of science.

Keywords: COVID-19 texts; evidentiality; interpersonal systems; popular science; specialized science

*Corresponding author: Jinjun Wang, School of Foreign Studies, Guangzhou University, Guangzhou, China, E-mail: jjwangkm@163.com
Jinyi Huang, Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies, Guangzhou, China, E-mail: brianhuang1988@163.com

1 Introduction

Evidentiality can be conceptualized, in a broad sense, as the speaker/writer's linguistic means of indicating the information source of the proposition (e.g. Aikhenvald 2004: 3–6; Boas 1938: 133; Mushin 2001: 9) making a commitment to the degree of its warrantability, credibility, and reliability (e.g. Boas et al. 1947: 206; Chafe 1986: 271; Sapir 1922: 144). It is a linguistic reflection of the nature of human consciousness that truth, reality, and fact are all relative, being subjectified and intersubjectified in human speech communities (Aikhenvald 2004: 218–219; Chafe and Nichols 1986: vii; Givón 1984). With a focus on evidentiality in scientific registers, several attempts have been made to describe the features of its different varieties: in textbooks and monographs on physics (Yang 2018); in academic journal articles on linguistics (Cui 2015; Yang 2009, 2015); in PhD dissertations on linguistics (Chen 2008); identifying the contrast between academic journal articles by English native writers and those by Chinese writers (Yang 2009, 2015); in review articles for academic purposes (Chen 2008, 2014); and identifying the contrast between specialized science texts and popular science texts in the disciplines of biology and biomedicine (Hidalgo-Downing 2017). However, the attention paid to evidentiality hitherto across a range of registerial variation in scientific texts is still far from sufficient, especially in mapping the variation between specialization and popularization in medical texts.

Recently, the appearance of the coronavirus COVID-19 and the global pandemic to which it has led have brought unprecedented challenges for countries around the world. In fighting the virus, the leading role of medical science has been a worldwide consensus, despite criticism from some anti-science voices. Under such conditions, the crucial function of evidentiality in medical science texts comes to the fore. "Science" itself is never a monolithic entity always understood in the same way, but rather a social construct produced by different groups with different interests (Hyland 2010: 118). In other words, the language of science can be taxonomized according to various semantic configurations as not only reshaping human experience ideationally (e.g. Halliday 2004 [1998a], 2004 [1998b]; Halliday and Martin 1993) but also enacting social relations interpersonally (e.g. Hyland 2005a, 2005b, 2009: 10–18, 2010). Such a taxonomy thus shapes "a continuum of discourse representation", with specialized science as the pole concerning "immanent value in specialized contexts" and popular science as focusing on "values […] to people's life" (Calsamiglia 2003: 140; de Oliveira and Pagano 2006; Pilkington 2018: 5), and with the whole continuum characterized by registerial variation according to different contextual configurations. In the interpersonal spectrum, the writer-reader relation is modified from that between peers in a shared domain of science to that

between popularizers and the lay public. Along this continuum, the text producer's lexicogrammatical and semantic choices in text instances vary accordingly, the meanings and wordings of evidentiality among them.

Hence, understanding evidentiality along the cline from specialization to popularization is an important aspect for unpacking the information provided in scientific texts. Based on the dataset of COVID-19 texts, this paper aims to respond to the questions listed below:

- Lexicogrammatically, what probabilistic patterns are realized by evidential usage in specialized science texts and popular science texts respectively?
- Semantically, how do the wordings in these two registers serve as semantic constructs for committing to modal responsibility and manipulating the dialogic space?

2 Theoretical framework

Answers to the question of the definition of evidentiality differ widely, including (i) whether it is grammatical or semantic; (ii) whether it focuses solely on information source or on a combination between information source and factual status; and (iii) whether it can be categorized as a subtype under the heading of "modality". Different answers to these questions have led to multiple classifications (e.g. Anderson 1986; Chafe and Nichols 1986; Jakobson 1957; Willett 1988). For the purpose of coping with the functions of evidentiality, not only in our collected data but also across a wider range of scientific texts, the theoretical framework here is set up by integrating previous classifications of evidentiality from both lexicogrammatical and semantic levels. Lexicogrammatically, Yang's (2009: 88, 2015: 62–66) EVIDENTIAL TAXONOMY and our own system of INFORMATION SOURCE are taken into consideration. For the former, Yang's evidentiality classification has been tested on a corpus of English academic discourse (approximately 670,000 words), and her functional perspective is more accessible for the current study. For the latter, INFORMATION SOURCE is designed to enable us to categorize different roles of the information source, essentially indicating who is responsible, in science and technology texts in general and medical science texts on COVID-19 in particular. Semantically, MODALIZATION (Halliday and Matthiessen 2014) and ENGAGEMENT (Martin and White 2005), whose combined descriptive scope fully covers our research object here, are set up as two parallel interpersonal systems, covering meanings of modal responsibility, i.e. to what extent the writer is responsible, and dialogic space, i.e. how does s/he react to his/her readership, respectively. From the MODALIZATION system, we have excerpted those subsystems that show the closest relation to evidentiality, i.e. Halliday and Matthiessen (2014: 182) ORIENTATION and

VALUE; while for ENGAGEMENT, we have generally followed Martin and White's (2005) description of ENGAGEMENT, choosing the features related to evidentiality in particular – Concurrence, Pronouncement, and Endorsement in dialogic contraction, and Entertaining, Acknowledgement, and Distancing in dialogic expansion.

The integration of these systems not only is convenient for text annotation and statistics using the UAM Corpus Tool 3.3 (abbr. UAMCT; O'Donnell 2013), but also helps to systemize evidentiality and examine its role as an instance of a "stylistic marker" (Aikhenvald 2004: 310–315; Arrese 2017: 195–223; Hidalgo-Downing 2017: 225–248). The whole design listed above is inputted into a "layer" panel in the UAMCT. The following sections will describe each component of EVIDENTIALITY in terms of its interpersonal strands at lexicogrammatical and semantic strata.

2.1 Lexicogrammar

In the lexicogrammar, evidentiality comprises two simultaneous systems – EVIDENTIAL TAXONOMY and INFORMATION SOURCE (Figure 1). This "simultaneity" presupposes that these two systems share the same entry condition (Eggins 2004: 195–198; Halliday and Matthiessen 2014: x; Matthiessen and Halliday 2009: 98).

In this sys-net, if the entry condition of "evidentiality-as-wording" is met, a language user's choices in the features of EVIDENTIAL TAXONOMY and INFORMATION SOURCE form a cross-classification. In other words, when analysts are tagging evidentiality at the lexicogrammatical stratum, they are required to co-select features in these two systems along a scale of increasing delicacy – from leftmost to rightmost in the sys-net (see example (1)). In all sentence examples, the information source is **_italicized and bolded_**; the evidential marker is _italicized and underlined_; and the propositional information is <u>double underlined</u>.

(1)   <u>Ensuring the stockpile is routinely monitored going forward must occur</u>, **_Schabacker_** _said_. (Dataset 2: _Modern Healthcare_, April 18, 2020)

In (1), evidentiality is lexicogrammatically realized by _Schabacker_ for INFORMATION SOURCE and _said_ for EVIDENTIAL TAXONOMY. As the delicacy moves rightwards in the sys-net, continuous choices in each of these two systems will show an increase in delicacy. For the choice of _Schabacker_, it assumes that the writer selects along an order of increasing delicacy in system of INFORMATION SOURCE: Information Source → Human → Specialist; for the selection of _said_, s/he makes choices along an analogous order in EVIDENTIAL TAXONOMY: Evidential Taxonomy → Reporting → Other-reporting → Verbal. In the following sections, we will provide syntagmatic characterizations of each of these two systems.
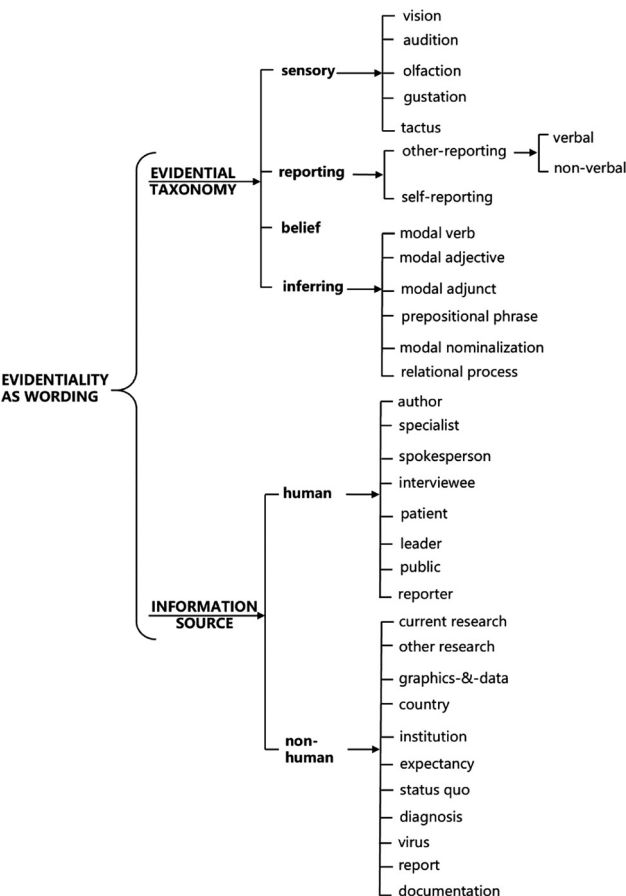
Figure 1: Sys-net of EVIDENTIALITY-AS-WORDING (adapted from Yang 2009, 2015).

2.1.1 Evidential taxonomy

EVIDENTIAL TAXONOMY is concerned with the following features (Table 1):

– **Sensory** is a feature that construes propositional information as first-hand experience obtained through human sensory channels such as Vision, Audition, Gustation, Olfaction, and Tactus.

– **Reporting** is a feature that identifies either an internal-oriented source (as in Self-reporting) or an external-oriented source (as in Other-reporting).

– **Inferring** is a feature conveying not only inferences by which deduction and induction are acquired from perceived and intuited evidence, but also

assumptions under which the inferential process is constructed on the basis of reasoning, assumptions, or general knowledge (Tang 2007: 68; Yang 2009: 91).

– **Belief** is a feature construing information in terms of the text organizer's mental state such as personal belief or opinion.

The typical realizations of these features are provided in Table 1 below. Each of the examples given is naturally occurring and retrieved from *Corpus of Contemporary American English* (abbr. *COCA*; Davies 2008).

2.1.2 Information source

INFORMATION SOURCE is a system representing the text producer's evaluation of the warrantability of the source of a proposition; which presupposes a paradox that the text organizer is in fact responsible for every single word in his/her text. However, in scientific writing, most of the modal responsibility supposed to be assigned to Authorship is conventionally shifted to other semantic roles. These semantic roles are defined by the second and third-order features of this system. As a proposition is commonly referred to a Human source, its information can be construed in terms of diverse positionings. The options in INFORMATION SOURCE cover but are not limited to the features listed in Table 2.

2.2 Semantics

At the semantic stratum, evidentiality is concerned with two aspects of interpersonal meaning: MODALIZATION and ENGAGEMENT (Figure 2).

In this sense, the interpersonal meanings of evidentiality show graduation in two ways: one is the degree to which the author is willing to show his/her responsibility for the reliability of information in the proposition; the other is the extent of the dialogic space the author opens up for the potential reader(s).

2.2.1 Modal responsibility

In terms of MODALIZATION (Halliday and Matthiessen 2014: 176–193), modal responsibility functions to construe how overtly the writer states his/her attitude towards the credibility of the given proposition (Thompson 2004: 67–75). In general, it can be realized as a textual deployment along the cline of subjectivity to objectivity (Table 3).

Table 1: Lexicogrammatical realizations of EVIDENTIAL TAXONOMY.

Evidential type	Feature		Typical realization
Sensory	**Vision**		*We see* that we have 5 independent variables or factors. (*COCA_ACAD* 2019)
			It has been seen that hypo-echoic plaque regions can be surrounded by the hyper-echoic regions. (*COCA_ACAD* 2016)
			See example data in 55D Fig (*COCA_ACAD* 2019)
	Audition		*We heard* that there was an opportunity to design an experiment that would go up on India's moonlander. (*COCA_MAG* 2017)
			It smells like someone is cooking potatoes. (*COCA_MAG* 1998)
	Olfaction		*We smell* food coming from your room. (*COCA_SPOK* 2017)
			It tastes like it's fresh. (*COCA_MAG* 2012)
	Gustation		*We taste* that particularly high calorie combination of sugar and fat. (*COCA_BLOG* 2012)
	Tactus		*My hand touched* something that hadn't been there an hour before. (*COCA_NEWS* 1990)
Reporting	**Other-reporting**	**Verbal**	*Chui says* that automation develops and is adopted slowly. (*COCA_MAG* 2019)
			As *is shown* in the work of the black churches and leaders. (*COCA_ACAD* 1997)
			It is said that someone has sky-colored eyes. (*COCA_ACAD* 2019)
		Non-verbal	That csnk-1 RNAI had no effect on the timing of centriole separation suggests that a force-independent licensing event is necessary to initiate separation *(Cabral et al. 2013 (9); Tsou and Stearns 2006 (70))*. (*COCA_ACAD* 2019)
			The framework reflects *a hypothesis* that the "capacity" of an organization to find and use research comprises the value placed on research evidence by the staff and organization. (*COCA_ACAD* 2019)
			According to **the court records**, James Ruman was charged with misdemeanor battery/making physical contact in 2008. (*COCA_NEWS* 2019)
			Eventually, we take into account trade openness (Chanda and Dalgaard 2008), measured through the ratio of trade to GDP (Figure 8). (*COCA_ACAD* 2019)
	Self-reporting		*We assume* that observations in set 1 satisfy the bivariate model. (*COCA_ACAD* 2019)

Table 1: (continued)

Evidential type	Feature	Typical realization
Belief		***I*** *think* that Donald Trump's campaign was a promise for near-term relief. (*COCA_ACAD* 2019)
		Professor Lyndel Prott *believes* that this interpretation of the article is too wide. (*COCA_ACAD* 2019)
Inferring	**Modal verb**	Pre-school-12 Education users *may* use this license as well. (*COCA_BLOG* 2012)
	Modal adjective	*It is possible* that the Indian hospital effluent could provide a particularly rich variety of carbapenemases. (*COCA_ACAD* 2019)
	Modal adjunct	Lots of people are *probably* going to lose their jobs. (*COCA_MAG* 2019)
	Relational process	*It seems* that the teaching of spelling is not always considered to be a priority. (*COCA_ACAD* 2019)
	Prepositional phrase	The Romantic era was, *to an extent*, a remake of the Heroic Age, minus most of the bloodshed. (*COCA_ACAD* 2008)
	Modal nominalization	To the best of our ability, we can only state *the possibility* that V2b-GABA exist. (*COCA_ACAD* 2019)

As the writer shows less of his/her subjectivity and more objectivity, so s/he takes less responsibility, and so the information contained in the proposition is more likely to be objectivized; while the writer shows more of his/her subjectivity and less objectivity, s/he takes more responsibility, and the propositional information is more likely to be subjectified. See Figure 3.

Therefore, there is cline from subjectivation to objectivation in terms of the degree of responsibility-taking.

2.2.2 Dialogic space

As defined by the system of ENGAGEMENT (Martin and White 2005), dialogic space serves to reveal the writer's manipulation of the possibilities of writer-reader dialogue, whereby the writer is able to show either solidarity or disalignment with putative readers. This system thus provides a global window through which we are able to identify the degree of negotiability/non-negotiability between writer and reader in texts. When the dialogic space is contracted, the authorial voice is given a relatively high status, leaving little or no space for negotiation for potential readers to express disagreement, challenge, or even aggression towards the given

Table 2: Lexicogrammatical realizations of INFORMATION SOURCE.

| Source type	Feature	Function	Typical realization
Human	Author	To indicate the proposition as an authorial production either explicitly or implicitly.	*I* think that <u>Donald Trump's campaign was a promise to near-term relief</u>. (*COCA_ACAD* 2019)
	Specialist	To frame the proposition as an authoritative, professional, and specialist construct.	***Professor Lyndel Prott*** believes that <u>this interpretation of the article is too wide</u>. (*COCA_ACAD* 2019)
	Spokesperson	To represent the proposition as a collective voice from a certain group sharing institutional interests.	<u>"On Wednesday, the group decided to mettings will stay private"</u>, ***the group's spokesman*** said. (*COCA_BLOG* 2012)
	Interviewee	To express the proposition as the record of a layperson's interviews.	***One interviewee*** said <u>"we have had customer feedback via Twitter and we've been able to act on it and show we listened"</u>. (*COCA_ACAD* 2012)
	Patient	To construe the proposition as the patient's response.	<u>"I feel like I did the best thing for the baby"</u>, ***the patient*** said. (*COCA_NEWS* 2019)
	Leader	To express the proposition as the statement of a Head of State.	***President Trump*** says <u>if he's impeached, the stock market will crash</u>. (*COCA_SPOK* 2019)
	Population	To express the proposition as the collective voice of the public.	If ***all the people*** believe <u>that his election was turning a page in history</u>, I think there will be this anxiety that we will move backwards. (*COCA_WEB* 2012)
	Reporter	To color the proposition as the voice of news media.	<u>He intends to appeal the conviction</u>, ***CNN*** reported. (*COCA_MAG* 2019)
Non-human	Current research	To express the proposition as a component of what the writer has found out in his/her research and writing.	***Our study*** highlights <u>that models fit to *in situ* water-quality data can be used to generate accurate predictions of TSS at both fresh-water and estuarine sites</u>. (*COCA_ACAD* 2019)
	Graphics-&-data	To construe the proposition as the content of statistic visualization.	***Figure 1*** shows <u>that for NAICS 447, the Midwest has consistently lower coverage</u>. (*COCA_ACAD* 2019)
	Other research	To turn the proposition into the findings of other research.	***Most studies*** have found <u>that these powers are mirroring leaders' influence over followers</u>. (*COCA_ACAD* 2018)

Table 2: (continued)

Source type	Feature	Function	Typical realization
	Country	To express the proposition as a country's official statement.	Only ***China*** stated <u>that it was in favor of the "total destruction" of these weapons</u>. (*COCA_ACAD* 1997)
	Status quo	To express the proposition as the results of current situation.	<u>Jimmy Reen was pronounced dead-strangled by the sheet he slept on</u>, according to ***the death certificate***. (*COCA_MAG* 1998)
	Diagnosis	To construe the proposition as something related to the diagnosis.	***Diagnosis*** requires <u>that there should be clear evidence of clinically significant impairment in social, academic, or occupational functioning</u>. (*COCA_WEB* 2012)
	Expectancy	To construe the proposition as something that has not happened.	<u>She attended a university that costs</u>, according to ***estimates***, <u>more than $70,000 per year</u>. (*COCA_MAG* 2012)
	Documentation	To construe the proposition as abiding by official documents or policies	<u>Tomas Dejesus was charged with possession with intent to deliver heroin</u>, according to ***documents***. (*COCA_WEB* 2012)
	Institution	To express the proposition as an institutional claim	<u>Influenza is recurring global health threat that</u>, according to the ***World Health Organization***, <u>is responsible for as many as 500,000 deaths</u>. (*COCA_MAG* 2017)
	Report	To express the proposition as part of an unofficial presentation of what has happened.	***The last report*** shows <u>unemployment is trying to sneak back up</u>. (*COCA_BLOG* 2012)

proposition; when the dialogic space is expanded, there are more possibilities for writer and reader to negotiate with each other (Table 4).
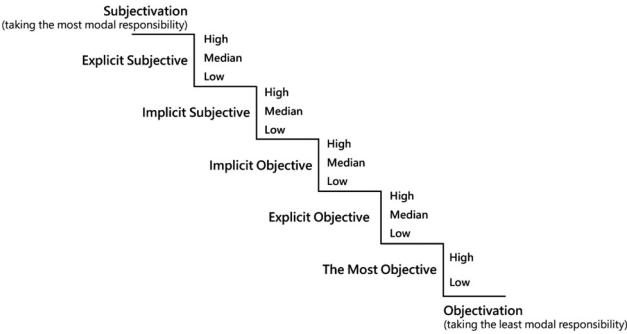
3 Methodology

3.1 Research method

In analyzing text as a semantic unit (Halliday 2002 [1981]), the "theory-driven" approach (Thompson and Hunston 2006: 1–14) of SFL enables us to pin down the
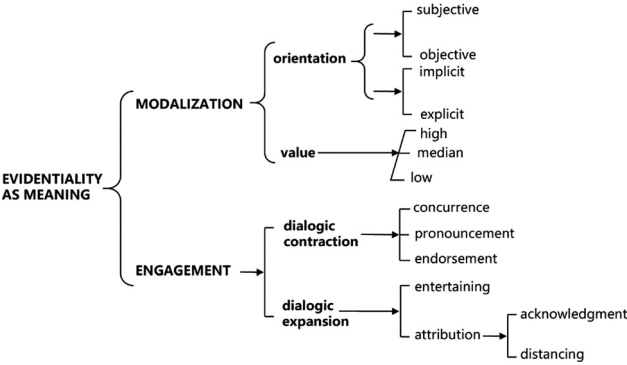
Figure 2: Sys-net of EVIDENTIALITY-AS-MEANING.[1]
(adapted from Halliday and Matthiessen 2014: 176–184; Martin and White 2005: Ch. 3).

Figure 3: The cline of objectivation.

semantic patterns in text (e.g. quantum of interaction, message, and flow of events) by the writer's particular choices and their relative frequency (Halliday 2005 [1991a]: 51–52, 2005 [1995]: 245–256; Halliday and Matthiessen 2014: 20–21, 659–665). The lexicogrammatical patterns are the realization of their semantic counterparts, and in written texts realize pretty much the whole of the semantics (Halliday and Matthiessen 2014: 659–660). In line with this characteristic, our research is put into practice by adopting "mixed-method research" (abbr. MMR) in

---

1 The way that we display the VALUE system is mainly based on Martin and White's (2005: 15–16) recognition of scaled systems. According to Martin and White (2005: 15–16), the paradigmatic representation of a scaled system is required to be formalized like a scale. Otherwise, there is no meaningful differentiation between subclassified system (such as ORIENTATION) and scaled system (such as VALUE).

Table 3: Evidential resources realizing MODALIZATION.

Orientation		Evidential type	Value		
			Low	**Median**	**High**
Implicit	**Subjective**	Inferring evidential	*may, might, can, could* *seem, appear*	*should, will, would*	*must*
Explicit	**Subjective**	Belief & Inferring evidential	*I/We believe it is possible that…*	*I/We believe it is likely that…*	*We believe it is…/ it is certain that…*
		Self-reporting evidential	*We/I discuss/suggest/say that…, etc.*		*We/I show/indicate that…, etc.*
Implicit	**Objective**	Inferring evidential	*possibly, maybe, perhaps, seemingly, apparently, etc.*	*Probably, etc.*	*certainly, indeed, evidently, obviously, clearly, etc.*
Explicit	**Objective**	Inferring evidential	*It is possible/ supposed/ plausible, etc. that…*	*It is likely that…*	*It is evident/clear/ certain/true etc. that…*
	The most objective	Other-reporting evidential	*other study argues that…, this analysis suggests that…, etc.*		*another study shows that…, etc.*

Table 4: Evidential resources realizing ENGAGEMENT.

Dialogism	Engagement feature	Evidential type	Examples
Contraction	**Proclaim: concur**	Inferring evidential	*must, of course, certainly, evidently, indeed,* etc.
	Proclaim: pronounce	Other-reporting evidential	(Authorial voice) *shows/demonstrates…,* etc.
	Proclaim: endorsement	Other-reporting evidential	*analysis shows/proves…, etc.*
Expansion	**Entertain**	Inferring evidential	*might, may, could, can, possibly, likely, supposed, maybe, perhaps, seemingly possibly, likely, supposed, maybe, perhaps, possibly, seemingly, etc.*
	Attribute: acknowledge	Self-reporting evidential	*We found, Our analysis shows…, etc.*
		Belief evidential	*We believe/expect/hope…, etc.*
		Other-reporting evidential	*X said/indicates/suggests…, Fact that…, argument that…, According to X, etc.*
	Attribute: distance	Other-reporting evidential	*X claim/allege/assert/reiterate…, etc.*

general, "QUAL→quan" progression[2] in particular (Riazi 2016: 189–193). In this sense, "qualitizing" (Nzabonimpa 2018) is our major concern, functioning not only to code the lexicogrammatical elements but also to interpret the meanings behind the data of related wordings; "quantitizing" (Nzabonimpa 2018) serves only to transform the coded features into numerical data, providing a readymade object for follow-up semantic interpretation.

Trinocularity in SFL – the views from which any linguistic phenomenon can be defined typically in terms of stratification (Matthiessen et al. 2010: 233–234) – enables this paper to adopt a bottom-up perspective, following a processing order from wording to meaning. Firstly, the lexicogrammatical realization of evidentiality in our collected texts will be identified semi-automatically by the UAMCT. Subsequently, the annotated results will be quantified in form of normalized frequencies and percentages whose significance will be tested automatically through Chi-Square test. Last but not least, we will interpret the statistics in the given data by considering its contextual cues.

3.2 Data collection

The current study takes its data from the discourse of both specialized science and popular science with the shared topic of COVID-19.

Firstly, this collection lays its emphasis on cross-register comparability. Interpersonally, the discourse of specialized science, characterized by communication between specialists in a homogeneous discourse community, aims to extend, transmit, and explore knowledge about natural science (Halliday 2004 [1988]). At the other pole, the discourse of popular science is concerned with a (re-) shaping process, initiated by experts, which aims to enlighten the lay readership in different communities in their understandings of scientific knowledge (Hyland 2010: 118). Registerial variation between the former and the latter entails a co-variance on both contextual and linguistic planes, with a change in the tenor of situational context from in-community dialogue to expert-layman transmission. In the case of our data, this shows up as differentiation in indicating the information source and committing to the reliability of the proposition. Under the circumstances of the coronavirus pandemic, to compare these two registers will further our understanding of how such co-variance occurs functionally.

---

2 This is based on the notation system developed to represent different MMR designs in terms of paradigm emphasis. Arrow signs (→) are used to represent sequential studies in which the qualitative tagging informs the instruments and procedures of quantitative data collection. The capital letters (i.e. QUAL) are used to show the paradigm focus (see Riazi 2016: 189–193).

Secondly, with a special focus on the topic of COVID-19, we must confirm the quality of texts included in our dataset so as to better fulfill our comparative needs. For this purpose, the texts of specialized science included come from the world's top four journals in medical science, *The Lancet*, *The Journal of American Medical Association*, *The British Medical Journal*, *The New England Journal of Medicine*. Those included in our corpus of popular science have been extracted from the popular science periodicals *Science Magazine*, *Modern Healthcare*,[3] *TechLife News*,[4] *New Scientist*,[5] and *Science Focus*.[6] The total number of words in each corpus is set up as a balancing criterion for this collection, with the dataset for specialized science comprising 16 texts amounting to 54,136 words, and that for popular science comprising 53 texts with 50,575 words.

Thirdly, validity and typicality for these data are guaranteed. The timeline for the included texts starts from January, 2020 – the time of the initial COVID-19 outbreak. After this, different text types on this topic began to be produced rapidly worldwide, and so the time range for the texts collected is from January, 2020 to April, 2020. In addition, it is worth acknowledging that each of the collected texts is officially published, which assumes a rigorous process of reviewing, editing, and fact-checking for each text.

---

3 *Modern Healthcare* is a "must-read" by the who's who in healthcare and has received many accolades listed in https://www.modernhealthcare.com/healthcare-editorial-awards-and-honors (accessed 15 May 2020). *Modern Healthcare* is the industry's leading source of healthcare business and policy news, research and information. It reports on important healthcare events and trends, as they happen, through its weekly print magazine, websites, e-newsletters, mobile products, and events.
4 *Techlife News* magazine offers up-to-the-minute coverage of the best new tech products, apps, and games alongside reviews.
5 The selection of *New Scientist* is mainly based on "Top 10 Science Magazine & Publication to follow in 2020" in https://blog.feedspot.com/science_magazines/ (accessed 15 May 2020). *New Scientist* is the world's most popular weekly science and technology magazine. It covers international news from a scientific standpoint, and asks the biggest-picture questions about life, the universe and what it means to be human.
6 The selection of *Science Focus (BBC)* is mainly based on "Top 10 Best Science Magazines in the World" in https://www.worldblaze.in/best-science-magazines/ (accessed 15 May 2020). *Science Focus (BBC)* is one of the best science magazines on the newsstand in the world now. The articles published in this magazine are basically based on recent innovations in science, all aspects of modern technology, space exploration, and gadgetry.

4 Data analysis and discussions

The output of our annotation results as a whole equals 3172 "feature usages"[7] (O'Donnell 2013: 33–36) in total – 1428 for the dataset of specialized science – abbreviated "Dataset 1" and 1744 for that of popular science – abbreviated "Dataset 2". Our investigation relies heavily on Halliday's (2005 [1991a]: 51, 2005 [1991b]: 64–70, 1996: 16–18, 2003: 23) notion of "probability", viewing the highly significant differences[8] between evidential choices in specialized science texts and those in popular science texts. In the following sections, a detailed analysis and description of these statistical divergences will be conducted in terms of EVIDENTIAL TAXONOMY, INFORMATION SOURCE, MODALIZATION, and ENGAGEMENT within our EVIDENTIALITY layer scheme.[9]

4.1 Lexicogrammar

4.1.1 Evidential taxonomy

As the probabilities in UAMCT indicate (Table 5), three of the four features can be seen as characterized by significant statistical differences. Reporting is selected more frequently in specialized science texts (1011/70.80%) than in popular science texts (964/55.28%). In Belief and Inferring, there is an opposite situation: the choices in popular science texts (36/2.06% for Belief; 712/40.83% for Inferring) outnumber those in specialized science texts (14/0.98% for Belief; 388/27.1% for Inferring). The above provides a statistical basis, for identifying patterns not only in the first-order features such as Sensory, Reporting, Belief and Inferring, but also their sub-features such as Other-reporting and Self-reporting in Reporting; Modal Verb, Modal Adjective, Modal Adjunct, Prepositional Phrase, Modal Nominalization in Inferring, etc. Noteworthily here, the probability outputted in UAMCT is divided into two

---

7 The notion of "feature usage" is a technical term used in UAMCT (O'Donnell 2013: 33–36), referred to as a specification of a feature in a layer (sys-net) and a program for describing the usage of features in the corpus at the layer (counts, mean, and standard deviation).
8 In UAMCT, the significance via Chi-Square test is graded into three levels (O'Donnell 2013: 34–36): high (+++), significance at the 98% level (2% chance of error); medium (++), significance at the 95% level (5% chance of error); weak (+), significance at the 90% level (10% chance of error). This paper is only concerned with high significance.
9 "Layer scheme" is, in UAMCT (O'Donnell 2013: 24), a technical term for describing the features the analyst wants to annotate the text with. Analyst has two options here: either to create new scheme or copy existing scheme. In this study, we choose to create a new layer scheme named EVIDENTIALITY.

visions: the global and the local (O'Donnell 2013: 35). The former represents the percentage as a proportion where a featured segment occupies a total number of coded features within an overall sys-net; the latter represents the percentage as a proportion where a featured segment occupies a total number of coded features in a particular system that includes the feature (O'Donnell 2013: 35).

Under the heading of Reporting, specialized science prefers Self-reporting (168/16.62%), while popular science favors Other-reporting (963/99.90%). See Table 6.

At this point, the tendency towards Self-reporting in specialized science texts gives us a hint that the propositional information in specialized science is semi-oticized as the researcher's possession – their own first-hand experience of the virus. In contrast, the statistical evidence that popular science tends more to Other-reporting shows that the proposition is more likely to be represented as a quotation by external voices, typically from specialists. The difference above can be exemplified by the different selections of information source, as shown in (2) and (3).

Table 5: Comparing the probability of evidential types: A global vision.

Feature	Specialized science		Popular science			
	Raw counts	Percentage	Raw counts	Percentage	Chi-square	Significance
Sensory	14	0.98%	31	1.78%	3.567	+
Reporting	1011	70.80%	964	55.28%	80.520	+++
Belief	14	0.98%	36	2.06%	5.945	+++
Inferring	388	27.17%	712	40.83%	64.625	+++
Total units	1428		1744			

Table 6: Comparing the probability in reporting: A local vision.

Feature	Specialized science		Popular science			
	Raw counts	Percentage	Raw counts	Percentage	Chi-square	Signif.
Other-reporting	843	83.33%	963	99.90%	171.976	+++
Self-reporting	168	16.62%	1	0.10%	171.976	+++
Total units	1011		964			
Verbal type of other-reporting	224	26.57%	899	93.35%	852.5	+++
Non-verbal type of other-reporting	619	73.43%	64	6.65%	852.5	+++
Total units	843		963			

(2)   ***We*** *recommend* that, in addition to using nucleic acid tests as the gold
        standard for the diagnosis of COVID-19 pneumonia, relevant clinical
        examinations are done, including blood counts and chest CT and a
        comprehensive evaluation of a patient's medical history, epidemiological
        exposure, and symptoms. (Dataset 1: *Lancet* 2020, 395: 809–815)

(3)   'I'm not going to judge people (who hook up two patients to a single
        ventilator) but the single fact of the matter is there is nothing as safe as'
        one ventilator for one patient, *said* ***Richard Branson, a respiratory
        therapist at University of Cincinnati College of Medicine and a
        member of the Society of Critical Care Medicine***. (Dataset 2: *Modern
        Healthcare* April 18, 2020)

(2) exemplifies the use of first-person pronoun as the information source, so that in
the context of Dataset 1 the information in the proposition can be construed as a
statement made by the writers themselves or specialists. For the Dataset 2, the
strategy shown in (3) is pervasive in borrowing specialists' professional voices to
state propositions that the writer supports, in order to enhance the warrantability
of the text and ultimately to gain a wider understanding from a lay readership. As
we move further into second-order feature in REPORTING, it is likewise observable that
popular science engages more in the verbal type of Other-reporting (as shown in
(3)) while specialized science prefers the non-verbal (see (4)).

(4)   To better reflect the dynamics of the COVID-19 epidemic and corresponding
        interventions, 5 periods were classified based on important dates that
        could affect the virus transmission in Wuhan ***(Figure 1)***. (Dataset 1: *JAMA*
        online April 10, 2020)

From the example above, we can sense that knowledge in specialized science is
construed not only through the open expression of *we* but also by the zero-form as
expressed by the passive form of the verb with no overt agent, guiding the putative
readers' attention to some external source of information (e.g. *Figure 1* as Graphics-
&-Data in (4)). Popular science, in contrast, shows a comparatively strong pref-
erence for nominalization (see (5)) and according-type (see (6)) within the non-
verbal type of Other-reporting.

(5)   Another method backed by *evidence* that it can clean N95 masks is
        vaporized hydrogen peroxide, which is also recommended by the CDC.
        (Dataset 2: *Modern Healthcare* April 18, 2020)

(6)   As a result, economic activity seems to be rebounding, *according to **a
        recent study from Imperial College London's COVID-19 Response Team***,
        of which Walter is a member. (Dataset 2: *New Scientist* April 11, 2020)

For the writer of popular science, evidential uses, such as examples (5) and (6) together with the verbal type in (3), demonstrates a commitment to linguistic credibility. By construing *it can be clean N95 masks* as an *evidence*, the case of (5) shows the writer's way of assigning the proposition to a Thing with factual status. It can otherwise be congruently reworded as *that it can be clean N95 masks is evident*. (6), in a non-verbal way, adopts an external voice, instead of authorial identity, as the information source for the proposition *economic activity seems to be rebounding*. These evidential usages, as a whole, appear to be dedicated to finding so-called "evidence". As our precondition for evidentiality states, the author of the text is supposed to be responsible for every single word, phrase, sentence, or citation occurred in his/her text, but in fact this is not the usual way. What is more interesting, as shown in (5) and (6), is that the writer's modal responsibility has been shifted in an implicit way.

As shown in Table 6, in popular science the choice of Belief is characterized by high local probability, while specialized science prefers the use of authorial voice by, for example, Self-reporting. Popular science emphasizes Belief, while specialized science attends to Self-reporting. While the writer's own voice may be presented explicitly in both registers, it is the trends in the evidential choices that make a difference. Furthermore, Self-reporting in specialized science prefers Material and Verbal Processes, characterizing the proposition as Goal and Verbiage respectively. In contrast, Belief in popular science centers on Mental Processes that can lead potential readers to treat the proposition as one of the author's subjective inner experiences such as imagination, consciousness, perception, etc. (Halliday and Matthiessen 2014: 251–253). See (7):

(7)   ***We*** don't yet <u>know</u> <u>if recovering from covid-19 makes you immune to the virus in the long term</u>. (Dataset 2: *New Scientist* April 11, 2020)

By adopting *we* as the information source and *know* as evidential marker, example (7) shapes a community between the writer and the reader as a collective Sensor and helps to construct the proposition as a mental experience that can be felt, thought of, and cognized. Understandably, the amount of such usage will be constrained, to a larger extent, in specialized science, because the conventional semiotic mode of specialized science does not allow subjective expressions to run over the whole text. The same also goes for popular science texts, yet at a comparatively lower level. Science specialization is typically valorized in terms of its discovery of nature and truth (Halliday and Martin 1993: 2–21), but populari-zation of science aims to recontextualize this immanent value of specialization by showing the consequence of such values and how their products are put to work in public lives (Calsamiglia 2003: 140). In this process of recontextualization, the

greater attention to Belief is one of the variational features that characterizes popular science.

Within the domain of Inferring, popular science prefers modal verbs (575/ 80.76%), whereas specialized science favors modal adjectives (64/16.49%) and modal adjuncts (80/20.62%). See Table 7.

Such patterns correspond with Halliday's (1985: 94–96, 2004 [1998a]: 43, 2004 [1998b]: 59–73) statements about differentiation in Mode, one of the three variables of situational context. As a text becomes more stylized, more technical and formal, the frequency of nominalization increases by re-construing more processes (i.e. verbs) into products/things (i.e. nouns). Conversely, as a text is characterized as more common-sense and informal, the frequency of verbalization goes up by transforming more human experiences into "doings" and "happenings" (i.e. verbs). In this sense, we argue that according to our comparison, specialized science is concerned with both nominalization and verbalization, but popular science places verbalization at more of a focal point. The pattern discussed above is also exemplified in (8), (9) and (10) below:

(8)   It is *notable* that healthcare workers as well as close contacts of previously confirmed patients were likely to have a good outcome […]. (Dataset 1: *BJM* 2020, 368, online published February, 19)

(9)   However, in three of these studies, the external validation datasets are *probably* not representative of the target population. (Dataset 1: *BJM* 2020, 369, online published April, 07)

(10)   The World Health Organization (WHO) currently only recommends face masks for people who are coughing or sneezing and for those who are caring for people who *may* have covid-19. (Dataset 2: *New Scientist* April 18, 2020)

Table 7: Comparing probability in the system of inferring: A local vision.

Feature	Specialized science		Popular science			
	Raw counts	Percentage	Raw counts	Percentage	ChiSqu	Signif.
Modal verb	235	60.57%	575	80.76%	52.742	+++
Modal adjective	64	16.49%	24	3.37%	58.773	+++
Modal adjunct	80	20.62%	86	12.08%	14.294	+++
Prepositional phrase	0	0.00%	0	0.00%	0	
Relational process	8	2.06%	27	3.79%	2.441	
Modal nominalization	1	0.26%	0	0.00%	1.837	
Total units	388		712			

The use of a modal verb is the most congruent form for evaluating the certainty of a proposition. In (10), the writer valorizes the credibility of the proposition *people have Covid-19* by a modal verb of lower certainty *may*. In this case, the authorial voice is masked by not referring to the author him/herself, for example, using personal pronouns such as *I*, *we*, etc. *May* here functions as a verbal operator realizing the Finite in the Mood structure, with the aim of maintaining the certainty of that proposition at a lower level and accordingly at a higher level its arguability value. Also serving the modal functions of verbalization is the modal adjunct, as shown in (9). In the syntagm, *probably* functions as modal adjunct within Mood structure, appraising the given proposition with a median value. The choice of *probably* implies that the possibility that *the external validation datasets are not representative of the target population* is neutral, leaving room for either extreme, i.e. positive or negative. Unlike (10) and (9), in (8) the evidential marker *notable* is not directly part of the Mood structure, but rather functions as the Head of a Complement in the Residue, attributing a higher probability both to the dummy subject (i.e. *it*) and the real subject (i.e. the whole proposition *healthcare workers as well as close contacts of previously confirmed patients were likely to have a good outcome* […]). In the Transitivity structure, a Relational Process (the attributive intensive type in particular) lies between the evidential marker *notable* and the given proposition. In other words, if the proposition is metaphorically nominalized as a Thing in the nominal group, the evidential marker *notable* can naturally function as Epithet modifying the values and qualities of the Thing.

4.1.2 Information source

The raw counts of all features in INFORMATION SOURCE are 3046. And they are distributed in terms of different preferences in two registers (Table 8).

Among the three first-order features in INFORMATION SOURCE, we see that the feature Human is the most frequent in the popular science texts (1348/83.06%), while Non-human (795/55.87%) and Unknown (32/2.25%) are more likely for the specialized science texts. As our focus moves towards higher delicacy into Human type, it can be seen that comparatively the specialized text's choices are much simpler – only Author (584/97.99%) and Patient (7/1.17%) – while the popular text's choices are multifarious: with Specialist (328/24.33%), Spokesperson (242/17.95%), Interviewee (105/7.79%), Leader (43/3.19%), and Reporter (21/1.56%). In making choices among Non-human features, the tendency towards monotony in the specialized text versus diversity in popular texts continues. The specialized texts feature a prominence of Graphics-&-Data (640/80.50%) and Current research (52/6.54%), whereas the popular texts foreground such features as Other research (44/16.18%),

Table 8: Comparing probability of choices at multiple delicacy in INFORMATION SOURCE: A local vision (with exclusion of insignificant results).

Feature	Specialized science		Popular science			
	Raw counts	Percent	Raw counts	Percent	ChiSqu	Signif.
Human	596	41.88%	1348	83.06%	556.666	+++
Non-human	795	55.87%	272	16.76%	509.556	+++
Unknown	32	2.25%	3	0.18%	28.436	+++
Total units	1423		1623			
Author	584	97.99%	601	44.58%	495.208	+++
Specialist	5	0.84%	328	24.33%	160.689	+++
Spokesperson	0	0.00%	242	17.95%	122.21	+++
Interviewee	0	0.00%	105	7.79%	49.075	+++
Patient	7	1.17%	4	0.30%	5.659	+++
Leader	0	0.00%	43	3.19%	19.442	+++
Reporter	0	0.00%	21	1.56%	9.386	+++
Total units of human	596		1348			
Other research	48	6.04%	44	16.18%	26.441	+++
Graphics-&-data	640	80.50%	7	2.57%	515.646	+++
Current research	52	6.54%	2	0.74%	14.216	+++
Country	7	0.88%	24	8.82%	45.326	+++
Institution	10	1.26%	130	47.79%	385.013	+++
Status quo	7	0.88%	8	2.94%	6.209	+++
Experience	11	1.38%	12	4.41%	8.811	+++
Virus	2	0.25%	5	1.84%	7.828	+++
Report	3	0.38%	11	4.04%	21.043	+++
Documentation	0	0.00%	20	7.35%	59.573	+++
Total units of non-human	795		272			

Country (24/8.82%), Institution (130/47.79%), Status quo (8/2.94%), Experience (12/4.41%), Virus (5/1.84%), Report (11/4.04%), and Documentation (20/7.35%).

Within specialized science texts, the interest in Non-human and Unknown is in tune with Prigogine and Stengers (1984: 6) standpoint that science, especially the natural sciences, as a successful dialogue with silent nature, often isolates humans from its object instead of bringing them closer to it, by revealing a dead, passive, and automated nature to humans. Rhetorically, it functions to construct the proposition as coming from an objective entity, with an obvious accommodation to the conventional criteria in scientific writing of clearness, objectivity, persuasiveness, and impersonality (Halliday 2004 [1997]: 181; Martin and Veel 1998).

In terms of Human source, the specialized texts pay more attention to the interaction between the scientists and their objects of study. Evidentiality is here deployed to construe propositional information either as diagnostic findings from

the writers – who are also experts in this field, as shown in example (2) above – or as the representation of patients' symptoms, conditions, reactions, etc. (see (11)).

(11)   Patients can find home isolation psychologically taxing, because ***they know*** that they are putting those they care about most at risk of contracting the disease. (Dataset 1: *Lancet* 2020, 395: 1305–14)

It is obvious that *know* in (11) is a belief evidential (i.e. Mental Process – Cognition type – in Transitivity structure) construing the proposition *they are putting those they care about most at risk of contracting the disease* as a Phenomenon sensed by *patients* as Sensor. The similar situation also goes for the reporting evidential through which the information is reported as Goal or Verbiage assigned modal responsibility by patients (Halliday and Matthiessen 2014: 225–228, 255). This suggests the writers' rigorous and detailed attitudes toward almost every aspect of the diagnosed patients.

 In order to evaluate the material, enhance its own persuasiveness, and acknowledge alternative views (Hyland 2005a: 66–67), specialized science symbolizes the collaboration between the writers' stance and external voices from other scholars' findings (Miller 1979: 616). Noteworthily, in terms of Non-human sources, the specialized texts seem however to be singing the opposite tune to the norm above by their non-prominent use of Other research and the prominent use of Graphics-&-Data and Current research. As a matter of fact, this is not a deviation from the norm but a unique rhetorical device in line with the classic belief of academic writing – representation of discourse about "Truth" (Lemke 1995: 178) and heavy dependence on fact-based reasoning, dispassionate observation, and the informed reflection. The citation of Graphics-&-Data functions to construe the source of the proposition as visualized data extracted from clinical results, and further to intensify the actuality and authenticity of the information provided. The citation of Current research (e.g. *our study*), mainly during the transition from data to findings, is adopted to elaborate on methodology and discussion which emphasize to some extent the subjectification of certain propositions.

 The ways scholars view the world are usually different from the laypeople's commonsense ways (Halliday 2004 [1995]: 7–8). In popular science, the transition between these two viewpoints tends to entail a re-selection of information source, for example, by construing its propositional information as specialists' findings, spokespersons' announcements, patients' symptoms, interviewees' on-site experiences, etc. For selections in Human and Non-human sources, two quantitative rankings can be listed to show that the choices in the popular texts are significantly more prominent than in the specialized texts: (i) in terms of Human, Specialist (328/24.33%) > Spokesperson (242/17.95%) > Interviewee (105/7.79%) > Leadership (43/3.19%) > Reporter (21/1.56%); (ii) in terms of

Non-human, Institution (130/47.79%) > Other research (44/16.18%) > Country (24/8.82%) > Documentation (20/7.35%) > Experience (12/4.41%) ≈ Report (11/4.04%) > Status quo (8/2.94%).

In terms of evidentiality, whether the popularization of science is successful seems to depend on its writers' strategic selections in both Human and Non-human options. Our data show that the specification of information source in popular science text is a process of value assignment to the given proposition. In the Human type, the proposition is most often construed as: (i) the voice of expertise that reinforces the professionalism and authoritativeness of its content, i.e. Expert; (ii) the collective stance of a group of people who share an institutional interest, i.e. Spokesperson; (iii) the voices of ordinary individuals that highlight their concerns about the experiences and opinions of members of society, i.e. Interviewee. Likewise, its Non-human counterparts also play a constructive role in the given propositions. The top priority is assigned to a construal identifying the propositional information as: (i) the outcome of collective decision-making in medical institutions, i.e. Institution; (ii) other medical scientists' research findings, i.e. Other research; and (iii) reports on the pandemic situation in particular countries, i.e. Country.

4.2 Semantics

4.2.1 Modal responsibility

"Modal responsibility" is a semantic concept that highlights the writer's position along a scale of how overtly s/he expresses his/her attitude towards the credibility of the propositions provided (Thompson 2004: 69–72). Statistical evidence for differences between the two registers is essential for explicating the different strategies of realizing modal responsibility. Simply in terms of ORIENTATION and VALUE, the difference is already clear to see: in VALUE, the specialized science texts favor High degree (377/48.90%) while the popular science texts prefer Low degree (974/68.88%); in ORIENTATION, the specialized texts favor Subjective (434/56.29%) and Explicit (221/28.66%), whereas the popular texts favor Objective (880/62.23%), especially The Most Objective (796/56.29%). However, the fact that this system involves two simultaneous subsystems with a shared entry condition (Halliday 1966; Tucker 1998: 40, 99–113; Eggins 2004: 195–197) necessitates further observation of the results of co-selections (Table 9).

Table 9 above demonstrates that the two registers differ at each point along the cline of objectivation. By foregrounding the different preferences in popular science and specialized science for cross-classified features in MODALIZATION, our

Table 9: Comparing the probability of co-selections in MODALIZATION.

Orientation	Corpus	Value												
		Low				Median				High			
		Raw	Percent	ChiSqu	Signif.	Raw	Percent	ChiSqu	Signif.	Raw	Percent	ChiSqu	Signif.
Implicit subjective	Set 1	159	64.63%	7.869	+++	83	33.73%	5.112	+++	4	1.62%	None	None
	Set 2	282	53.61%			224	42.58%			10	1.90%		
Explicit subjective	Set 1	66	36.66%	18.967	+++	0	0.00%	None	None	114	63.33%	18.967	+++
	Set 2	14	100%			0	0.00%			0	0.00%		
Implicit objective	Set 1	20	23.52%	39.636	+++	18	21.17%	7.985	+++	47	55.29%	18.709	+++
	Set 2	59	73.75%			4	5.00%			17	21.25%		
Explicit objective	Set 1	15	36.58%	8.411	+++	22	53.65%	5.772	+++	4	9.75%	None	None
	Set 2	8	100%			0	0.00%			0	0.00%		
The most objective	Set 1	11	5.21%	359.648	+++					200	94.78%	359.648	+++
	Set 2	609	76.89%							183	23.10%		

comparative interpretation has been organized along the cline of "the extent to which the writer openly accepts responsibility for the subjective assessments being expressed" (Thompson 2004: 71).

Along this cline, specialized science text favors each value – Low value (incl. Implicit Subjective), Median value (incl. Implicit Objective, Explicit Objective), and High value (incl. Explicit Subjective, Implicit Objective, and The Most Objective). With Explicit Subjective, the proposition is construed more often as an assertive and proclamatory statement for which the specialized writer as information source is explicitly responsible. The writer also frequently uses the low-value Implicit Subjective frequently to shift responsibility, building an alignment with alternative stances and leaving space for further dialogue and negotiation. In Objective domains, the value returns to Median and High: by the choice of Implicit Objective and Explicit Objective, the concealed identity of authorship evaluates the propositions as comparatively reliable, credible, and warrantable, without taking the risk of revealing their own voices; at the highest level of objectivation (i.e. The Most Objective), the proposition is construed as information either from outside sources or from the findings of current research. At the same time, evidentials for valorizing the factual status of information help to add high certainty and reliability to propositions. There is a tendency towards taking modal responsibility in specialized science texts: as the position moves further toward the objective end along the cline, more High value choices are made.

Specialized science texts, as a channel for communicating the authoritativeness of science, regard COVID-19 as not purely an academic issue but as a material threat that endangers all humankind. If most propositions in texts are construed as uncertain and unknown, the power of science will be overshadowed, and this is likely to spread the global panic. To avoid these possibilities, most propositions given in relevant research are construed as certain and assured, for the most part leaving the uncertain and the unassured unsaid. This corresponds to our statistical evidence that propositions deriving from academic community voices tend to be valorized as high, and in fulfilling this need, evidentiality plays an important role.

The popular texts show a different weighting to the specialized texts. In general, they tend to highlight a Low value of Explicit Subjective, Implicit Objective, Explicit Objective, and The Most Objective and a Median value of Implicit Subjective. By the choice of Explicit Subjective, propositions are frequently construed as negotiable and discussable objects overtly originating from the popularizer's subjectivity. By making subjectivity implicit, the value increases to Median, underlining Hyland's (2005a: 99) conceptualization of "newsworthiness" – news value of propositions without their certainty or uncertainty being overstated. In the domains of Objectivity, including Implicit Objective, Explicit Objective, and The Most Objective, propositions are consistently valorized as Low.

To wrap up, the popularizer takes less modal responsibility, by contrast with the specialist.

Unlike the specialized texts that focus on the whole procedure of research including context, methods, discussion, findings, etc., the popularizer deals more with disseminating rather than pondering on the research findings. Since one of its keys is "entertainment" (Hyland 2009: 154), popular science functions to communicate scientific information about COVID-19 to a wider range of audiences at a more easily accessible level. As a linguistic reflection of the growing impact of science on the lives of the general public, it aims to stimulate lay readers' interest in science and is willing to embrace a wide range of controversies and debates on science-related topics. In such dialogic circumstances, propositions are naturally construed as negotiable and open-ended. The reality, actuality, and truthfulness included in these propositions can bear any form of potential examination, revision, and update.

4.2.2 Dialogic space

That evidentiality functions to manipulate the dialogic space requires relevant statistical evidence. To this end, UAMCT allows us to automate a data output of the coded features in both datasets (Table 10). Due to the statistical insignificance of the results for Acknowledgement and Distancing, the data for Attribution in Table 10 below will not be specified as to its sub-features. Comparing the two datasets at first-order delicacy, specialized science texts favor dialogic expansion (355/44.60%) while popular science texts prefer dialogic contraction (1118/77.37%). When focusing on the resource of dialogic expansion, for specialized science texts it is foregrounded in Entertaining (326/73.92%), whereas for popular science texts it is foregrounded in Attribution (772/69.05%). When we examine the resources of Contraction, popular science texts show a tendency to Concurrence (259/79.20%), while specialized science texts lean to Endorsement (125/35.21%) and Pronouncement (110/30.99%).

In specialized science texts, the overall tendency of evidential usage towards dialogic expansion implies the specialists' strong inclination to create an image of solidarity and thus make allowance for positions that are dialogically alternative to the authorial position. In doing so, most authorial positions are stated as one possibility among a range of potential stances available to putative readers. From the resource of dialogic expansion, the specialists choose Entertaining to state that the possibility of alternative positions co-exists with theirs. Such a usage functions to lower the authority of authorial voice in expressing propositions in the specialized texts. Usually, the author's voices are silenced, and their implicitness is used to advocate dialogic inclusiveness where their voices or the other alternatives

Table 10: Comparing probability of selections in ENGAGEMENT: A local vision.

Corpus	Dialogic expansion		Dialogic contraction	
	Raw counts	Percentage	Raw counts	Percentage
Set 1	355	44.60%	441	55.40%
Set 2	327	22.63%	1118	77.37%
ChiSqu	116.999		116.999	
Signif.	+++		+++	

Corpus	Entertain		Attribute		Proclaim					
					Concur		Endorse		Pronounce	
	Raw	Percent	Raw	Percent	Raw	Percent	Raw	Percent	Raw	Percent
Set 1	326	73.92%	115	26.08%	120	33.80%	125	35.21%	110	30.99%
Set 2	346	30.95%	772	69.05%	259	79.20%	65	19.88%	3	0.92%
ChiSqu	238.156		238.156		142.112		111.325		19.913	
Signif.	+++		+++		+++		+++		+++	

are treated only as one of the possible choices. Conventionally, interpersonal rhetoric in the specialized science texts is never characterized as peremptory, dogmatical or arbitrary but as inclusive, expansive, and even comprehensive.

The specialized science texts also tend to contract the dialogic scope, though their expansive choices are more prominent. Dialogic contraction requires Endorsement to construe the proposition as other researchers' findings, functioning not only to conceal their own attitudes and responsibilities but to maintain their position as authority in an only implicit way. In a "non-self" way, the validity, facticity, correctness and sometimes even undeniability of their propositions are maximized without any risk of exposing their own voices. In addition, Pronouncement, as a dialogically contractive resource that explicitly intensifies, emphasizes, and interpolates the authorial voice, is also selected to clarify the degree to which the factual status of proposition is highlighted and authorial subjectivity engaged. In this sense, the evidentials that function to realize Pronouncement are adopted to contract the dialogic space, but still leave room for authorship to be expressed in a covert way by masking his/her subjectivity.

On the other hand, the popular science texts' preference for dialogic contraction suggests their aim of establishing an authoritative identity *vis-à-vis* lay readers. Most of their evidential choices function to magnify the authorial voice, in order to keep off, limit, and even challenge alternatives. In so doing, the popularizer's position of power over lay people can be stabilized. By using Concurrence, propositional information is more often construed as an assured consensus and

sometimes even as a "promise". Its validity, tenability, and reliability can thus be construed as a multilateral recognition. In this sense, contracting the dialogic scope signifies a focus solely on what has been reported in the proposition to build an alignment solely with positive evaluations such as acceptance, acknowledgement, agreement, etc. that are shared by author and potential readers.

In realizing dialogic expansion, Attribution is adopted to engage in a potential negotiation between the authorial voice and the alternatives. To quote external voices as information source, mostly Specialist, Spokesperson, or Interviewee, and to adopt the neutral-value evidentials, mostly *say*, together function to shift the popularizer's responsibility and leave their own positioning unsaid.

5 Conclusion

The special period of the COVID-19 pandemic has presented more urgent demands for the evidence-based dissemination of knowledge about medical science, especially epidemiology. This paper seeks for a co-variance between evidentiality-as-wording and evidentiality-as-meaning in two registers – specialized science and popular science. On the lexicogrammatical stratum, evidentiality in both registers is realized as different distribution patterns of EVIDENTIAL TAXONOMY and INFORMATION SOURCE. In terms of EVIDENTIAL TAXONOMY, specialized texts stress the importance of Reporting, while popular texts attach more importance to Belief and Inferring. In terms of INFORMATION SOURCE, specialized texts feature Author and Patient in Human source, and Graphics-&-Data and Current research in Non-human source; popular texts feature Specialist, Spokesperson, and Interviewee in the Human, and Other research, Country, and Institution in the Non-human. Based on this, it can be concluded that the lexicogrammatical realizations between the two sets are divergent in their ways of knowledge building. Specialized science texts highlight the reports of what the writers and their research have found out and how patients have reacted and responded. In contrast, the popular science texts underline the state of mind among scientists, countries, institutions, and civilians, and the degree to which authorship is hidden and the responsibility for the factual status of the proposition is shifted to other information source.

At the semantic stratum, those wordings of evidentiality function to generate interpersonal meanings of taking/shifting modal responsibility and manipulating dialogic space. Under MODALIZATION and ENGAGEMENT, a clear differentiation is also seen: in terms of modal responsibility, specialists take on more than do the popularizers; in terms of dialogic scope, specialists are more interested in expansion (mainly through Entertaining) but popularizers in contraction (mainly through Concurrence). This leads us to the conclusion that (i) knowledge in specialized

science is built up in such a way that the writer takes on more modal responsibility and, at the same time, leaves openness to alternative values; while (ii) knowledge building in popular science is reached by shifting the modal responsibility and accentuating the writer's powerful identity as a possessor of scientific knowledge.

With COVID-19 texts as a case study, the lexicogrammatical and semantic variations of evidential use have provided evidence of the divergences between the texts of specialized science and those of popular science in the medical realm. Such textual differentiation should not be limited to the topic of medical science. More importantly, it can be dedicated to setting up "evidentiality" as a text-based approach to discriminating between specialization and popularization in all sciences.

References

Aikhenvald, Alexandra Yurievna. 2004. *Evidentiality*. Oxford: Oxford University Press.

Anderson, Lloyd B. 1986. Evidentiality, path of change, and mental maps: Typologically regular asymmetries. In Wallace Chafe & Johanna Nichols (eds.), *Evidentiality: The linguistic coding of epistemology*, 273–312. Norwood, NJ: Albex.

Arrese, Juana Isabel Marín. 2017. Multifunctionality of evidential expressions in discourse domains and genres: Evidence from cross-linguistic case studies. In Juana Isabel Marín Arresse, Gerda Haßler & Marta Carretero (eds.), *Evidentiality revisited: Cognitive grammar, functional and discourse-pragmatic perspectives*, 195–224. Amsterdam & Philadelphia: John Benjamins.

Boas, Franz. 1938. Language. In Franz Boas (ed.), *General anthropology*, 124–145. Boston: D. C. Heath and Company.

Boas, Franz, Helene Boas Yampolsky & Zellig Sabbettai Harris. 1947. Kwakiutl grammar with a glossary of the suffixes. *Transactions of the American Philosophical Society* 37(3). 201–377.

Calsamiglia, Helena. 2003. Popularization discourse. *Discourse Studies* 5(2). 139–146.

Chafe, Wallace. 1986. Evidentiality in English conversation and academic writing. In Wallace Chafe & Johanna Nichols (eds.), *Evidentiality: The linguistic coding of epistemology*, 261–272. Norwood, NJ: Albex.

Chafe, Wallace & Johanna Nichols. 1986. *Evidentiality: The linguistic coding of epistemology*. Norwood, NJ: Ablex.

Chen, Mingfang. 2008. *Xueshu yupian de zhuanshu he shijiao: Jiyu yingyu boshi xuewei lunwen wenxian zongshu de yanjiu* [*Reporting and perspectives in academic discourse: A study on literature reviews of English dissertations*]. Beijing: Foreign Language Teaching and Research Press.

Chen, Zheng. 2014. *Jiyu zhuguaxing he jiaohu zhuguanxing lianxutong de yupian yanjuxing fenxi: yi lunbian xing yupian weili* [*Evidentiality in discourse: A corpus-based study of English argumentative discourse*]. Shanghai: Shanghai International University dissertation.

Cui, Lin. 2015. *Renzhi yuyong shijiao xia yingyu xueshu lunwen zhong de yanju biaoji yanjiu* [*A cognitive-pragmatic study of evidential markers in English academic papers*]. Changchun: Northeast Normal University dissertation.

Davies, Mark. 2008. The corpus of contemporary American English (COCA): 520 million words, 1990–present. http://corpus.byu.edu/coca/ (accessed 15 January 2021).

de Oliveira, Janaina Minelli & Adriana Silvina Pagano. 2006. The research article and the science popularization article: A probabilistic functional grammar perspective on direct discourse representation. *Discourse Studies* 8(5). 627–646.

Eggins, Suzanne. 2004. *An introduction to systemic functional linguistics*, 2nd edn. London & New York: Bloomsbury.

Givón, Thomas. 1984. Evidentiality and epistemic space. *Studies in Languages* 6(1). 23–49.

Halliday, Michael Alexander Kirkwood. 1966. Some notes on "deep" grammar. *Journal of Linguistics* 2(1). 57–67.

Halliday, Michael Alexander Kirkwood. 1985. *Spoken and written language*. Oxford & New York: Oxford University Press.

Halliday, Michael Alexander Kirkwood. 1996. On grammar and grammatics. In Ruqaiya Hasan, Camel Cloran & David G. Butt (eds.), *Functional description: Theory in practice*, 1–38. Amsterdam: John Benjamins.

Halliday, Michael Alexander Kirkwood. 2002 [1981]. Text semantics and clause grammar: How is a text like a clause? In Jonathan James Webster (ed.), *On grammar: Vol. 1 in the collected works of M. A. K. Halliday*, 219–260. London & New York: Continuum.

Halliday, Michael Alexander Kirkwood. 2003. Introduction: On the "architecture" of human language. In Jonathan James Webster (ed.), *On language and linguistics: Vol. 3 in the collected works of M. A. K. Halliday*, 1–29. London & New York: Continuum.

Halliday, Michael Alexander Kirkwood. 2004 [1988]. On the language of physical science. In Jonathan James Webster (ed.), *The language of science: Vol. 5 in the collected works of M. A. K. Halliday*, 140–158. London & New York: Continuum.

Halliday, Michael Alexander Kirkwood. 2004 [1995]. Language and the reshaping of human experience. In Jonathan James Webster (ed.), *The language of science: Vol. 5 in the collected works of M. A.K. Halliday*, 7–23. London & New York: Continuum.

Halliday, Michael Alexander Kirkwood. 2004 [1997]. On the grammar of scientific English. In Jonathan James Webster (ed.), *The language of science: Vol. 5 in the collected works of M. A. K. Halliday*, 181–198. London & New York: Continuum.

Halliday, Michael Alexander Kirkwood. 2004 [1998a]. Language and knowledge: The "unpacking" of text. In Jonathan James Webster (ed.), *The language of science: Vol. 5 in the collected works of M. A. K. Halliday*, 24–48. London & New York: Continuum.

Halliday, Michael Alexander Kirwood. 2004 [1998b]. Things and relations: regrammaticizing experience as technical knowledge. In Jonathan James Webster (ed.), *The language of science: Vol. 5 in the collected works of M. A. K. Halliday*, 49–101. London & New York: Continuum.

Halliday, Michael Alexander Kirkwood. 2005 [1991a]. Towards probabilistic interpretations. In Jonathan James Webster (ed.), *Computational and quantitative studies: Vol. 6 in the collected works of M. A. K. Halliday*, 42–62. London & New York: Continuum.

Halliday, Michael Alexander Kirkwood. 2005 [1991b]. Corpus studies and probabilistic grammar. In Jonathan James Webster (ed.), *Computational and quantitative studies: Vol. 6 in the collected works of M. A. K. Halliday*, 63–75. London & New York: Continuum.

Halliday, Michael Alexander Kirkwood. 2005 [1995]. Computing meanings: Some reflections on past experience and present prospects. In Jonathan James Webster (ed.), *Computational and quantitative studies: Vol. 6 in the collected works of M. A. K. Halliday*, 239–267. London & New York: Continuum.

Halliday, Michael Alexander Kirkwood & Christian Matthias Ingemar Martin Matthiessen. 2014. *Halliday's introduction to functional grammar*, 4th edn. London & New York: Routledge.

Halliday, Michael Alexander Kirkwood & James Robert Martin. 1993. *Writing science: Literacy and discursive power*. London & Washington, DC: Falmer.

Hidalgo-Downing, Laura. 2017. Evidential and epistemic stance strategies in scientific communication: A corpus study of semi-formal and expert publications. In Juana Isabel Marín Arresse, Gerda Haßler & Marta Carretero (eds.), *Evidentiality revisited: Cognitive grammar, functional and discourse-pragmatic perspectives*, 225–248. Amsterdam & Philadelphia: John Benjamins.

Hyland, Ken. 2005a. *Metadiscourse*. London & New York: Continuum.

Hyland, Ken. 2005b. Stance and engagement: A model of interaction in academic discourse. *Discourse Studies* 6(2). 173–191.

Hyland, Ken. 2009. *Academic discourse: English in a global context*. London & New York: Continuum.

Hyland, Ken. 2010. Constructing proximity: Relating to readers in popular and professional science. *Journal of English for Academic Purposes* 9. 116–127.

Jakobson, Roman. 1957. *Shifters, verbal categories, and the Russian verb*. Harvard University: Department of Slavic Languages and Literatures.

Lemke, Jay L. 1995. *Textual politics: Discourse and social dynamics*. London & Bristol, PA: Taylor & Francis.

Martin, James Robert & Peter Robert Rupert White. 2005. *The language of evaluation: Appraisal in English*. London & New York: Palgrave Macmillan.

Martin, James Robert & Robert Veel. 1998. *Reading science: Critical and functional perspectives on discourse of science*. London: Routledge.

Matthiessen, Christian Matthias Ingemar Martin & Michael Alexander Kirkwood Halliday. 2009. *Systemic functional grammar: A first step into the theory*. Beijing: Higher Education Press.

Matthiessen, Christian Matthias Ingemar Martin, Kazuhiro Teruya & Marvin Lam. 2010. *Key terms in systemic functional linguistics*. London & New York: Continuum.

Miller, Carolyn Rae. 1979. A humanistic rationale for technical writing. *College English* 40(6). 610–617.

Mushin, Ilana. 2001. *Evidentiality and epistemological stance: Narrative retelling*. Amsterdam & Philadelphia: John Benjamins.

Nzabonimpa, Jean Providence. 2018. Quantitizing and qualitizing (im-)possibilities in mixed methods research. *Methodological Innovations* 11(2). 1–16.

O'Donnell, Michael. 2013. UAM corpus tool manual. http://www.corpustool.com/Documentation/UAMCorpusToolTutorial3.0.pdf (accessed 28 April 2020).

Pilkington, Olga. 2018. *Presented discourse in popular science: Professional voices in books for lay audiences*. Leiden & Boston: Brill.

Prigogine, Ilya & Isabelle Stengers. 1984. *Order out of chaos: Man's new dialogue with nature*. London: Heinemann.

Riazi, Mehdi. 2016. *The Routledge encyclopedia of research methods in applied linguistics: Quantitative, qualitative, and mixed-methods research*. London & New York: Routledge.

Sapir, Edward. 1922. *Language: An introduction to the study of speech*. New York: Harcourt, Brace and Company.

Tang, Bin. 2007. *Yingyu yiqing xinwen zhong yanjuxing yupian tezheng de xitong gongneng yanjiu* [*Systemic functional approach to discourse features of evidentiality in English news reports of epidemic situation update*]. Shanghai: Fudan University dissertation.

Thompson, Geoff. 2004. *Introducing functional grammar*, 2nd edn. London: Hodder Education.

Thompson, Geoff & Susan Hunston. 2006. *System and corpus: Exploring connections*. London & New York: Equinox.

Tucker, Gordon Howard. 1998. *The lexicogrammar of adjectives: A systemic functional approach to lexis*. London & New York: Cassell.

Willett, Thomas. 1988. A crosslinguistic survey of the grammaticization of evidentiality. *Studies in Language* 12. 51–97.

Yang, Linxiu. 2009. *Yingyu keyan lunwen zhong de yanjuxing* [*Evidentiality in English research articles*]. Xiamen: Xiamen University dissertation.

Yang, Linxiu. 2015. *Yingyu xueshu lunwen zhong de yanjuxing yanjiu* [*Evidentiality in English linguistics research articles*]. Beijing: Science Press.

Yang, Xinzhang. 2018. Yingyu wulixue yupian zhong de yanjuxing dongci [Evidential verbs in English physics texts]. *Yingyu Yanjiu* [*English Studies*] (1). 66–76.