

Making Respiratory Medicine Visible: Intralingual and Multimodal Mediation in Academic and Institutional Health Videos

Michela Canepari

Department of Humanities, Social Sciences and Cultural Industries, University of Parma

Via D'Azeglio 85, Parma, Italy

E-mail: michela.canepari@unipr.it

ORCID: orcid.org/0000-0002-7285-8503

Received: September 11, 2026

Accepted: October 1, 2026

Published: October 8, 2026

doi:10.5296/ijl.v18i5.24079

URL: <https://doi.org/10.5296/ijl.v18i5.24079>

Abstract

Audiovisual health communication must preserve terminological accuracy while making complex physiological processes intelligible to viewers with different levels of medical knowledge. This article presents a comparative multimodal case study of four English-language audiovisual resources produced or hosted in British healthcare and academic contexts, organised into two academic-institutional pairs: the International Primary Care Respiratory Group's *How We Breathe* and the NHS animated video *Asthma*; and Peter Barnes's HSTalks lecture *COPD and Asthma: Similarities and Differences* and Blackpool Teaching Hospitals' *Symptoms of COPD*. The analysis examines how spoken explanation, curated intralingual captions or subtitles, on-screen written labels, animation, slides, emphatic gestures, temporal organisation, and audience positioning jointly mediate respiratory knowledge. It combines close linguistic analysis with speech-caption, temporal, and visual annotation. The findings suggest that accessibility in these resources is constructed primarily through strategies other than the removal of specialist terminology. The resources deploy different combinations of dual-register mediation, selective terminological retention, contrastive visual architecture, causal compression, segmentation, and cross-modal access. The comparison thus distinguishes information reduction from lexical simplification and proposes mediation density and causal resolution as qualitative parameters for analysing audiovisual medical accessibility. On this basis, the study advances a context-sensitive account of accessibility grounded in the distribution of explanatory work across modes. It also demonstrates the value of treating intralingual captions as one component of a wider

multimodal communicative design rather than as the sole site of intralingual translation.

Keywords: Health communication, Medical English, Audiovisual translation, Accessibility, Intralingual captions, Multimodality

1. Introduction

Increasingly, medical knowledge circulates through publicly accessible audiovisual media. Academic organisations, professional associations, and healthcare institutions use videos to explain anatomy, pathophysiology, symptoms, and treatment to audiences whose linguistic and disciplinary knowledge varies considerably. These materials therefore face a persistent communicative problem: they must retain enough technical precision to represent medical knowledge accurately while enabling non-specialist viewers to understand processes that are abstract, dynamic, and often invisible.

This problem is particularly evident in respiratory communication. Terms such as “ventilation”, “respiration”, “bronchospasm”, “bronchoconstriction” and “atopy” refer to clinically meaningful distinctions, but verbal definitions alone may not enable viewers to construct a usable mental model of breathing or airway obstruction. Audiovisual communication can thus distribute explanatory work across speech, captions, written labels and animation, so that accessibility may arise without systematic replacement of medical terminology by simpler vocabulary.

Intralingual translation naturally provides a productive perspective on this process because it encompasses reformulation within the same language in response to changes in audience, purpose, register and knowledge. In audiovisual texts, however, intralingual mediation cannot be located only in the relationship between speech and captions: the spoken discourse may in fact already contain definitions, analogies and lay equivalents; visual labels may retain technical terminology; and animation may explain spatial or causal relations that neither verbal channel states independently. The relevant unit is therefore the coordinated multimodal event.

The present article reports an intensive multimodal analysis of four purposively selected health videos organised into two academic–institutional pairs. The first, referred to here as the breathing–asthma pair, comprises the International Primary Care Respiratory Group’s *How We Breathe* (2023) and the NHS video *Asthma* (2024). The former develops a broad account of breathing, breathlessness, and respiratory intervention and explicitly seeks to improve ‘breathing literacy’, whereas the latter offers a short, disease-specific explanation of asthma and its management.

The COPD pair comprises Peter Barnes’s HSTalks lecture *COPD and Asthma: Similarities and Differences* (2021), produced within an academic biomedical context, and Blackpool Teaching Hospitals NHS Foundation Trust’s *Symptoms of COPD* (2026). Read together with the first pair, these resources allow the same analytical dimensions to be examined across different configurations of expert explanation, animation, slides, images, written text, emphatic gestures, and intralingual subtitles and captions. (Note 1) Although the comparison does not include patient testimony, it examines how accessibility is constructed when

experiential voice is absent and when human presence serves different epistemic and interpersonal functions.

The two pairs are not treated as perfectly equivalent accounts of identical content. The academic resources primarily address professionals, trainees or mixed educational audiences, whereas the institutional resources address patients, carers and the lay public. Their functional relationship nevertheless permits comparison between elaborative educational models and compressed institutional ones, while the within-pair organisation limits claims based on unequal duration and scope.

The study addresses two research questions: 1. How do the four resources coordinate specialist and lay language, captions, written text and visual representation to mediate respiratory knowledge? 2. How do differences in communicative purpose and modal configuration affect mediation density, causal resolution and the construction of cognitive, sensory, affective and operational accessibility?

The analysis therefore extends work on intralingual translation and audiovisual accessibility by treating accessibility as a context-sensitive effect of linguistic, temporal, and intersemiotic organisation. Its contribution lies in distinguishing lexical simplification from terminological pairing, selective scope, causal compression, embodied explanation, cross-modal access, and distributed multimodal explication.

2. Theoretical Orientation

2.1 Intralingual Translation and Specialised Discourse

Jakobson's (1959) account of intralingual translation as rewording established a broad category for interpreting verbal signs through other signs of the same language. Later research has shown that intralingual translation may involve changes in register, knowledge level, communicative purpose and receiver profile rather than simple synonymic substitution (Korning Zethsen, 2009, 2021; Korning Zethsen & Hill-Madsen, 2016; Canepari, 2023, 2024, forthcoming). Although intralingual translation does not necessarily entail simplification or the pursuit of greater accessibility (Canepari, 2024), studies of specialised discourse (Hill-Madsen, 2015, 2019, 2022; Peruzzo & Rocco, 2022) have examined strategies such as explication, reformulation, reduction, and lexical adjustment as potential means of enhancing accessibility.

In medical communication in particular, these strategies negotiate the tension between comprehensibility and accuracy. While technical terms may obstruct access for non-specialists, their removal is not always desirable, as it may compromise the precision or completeness of the information conveyed. Retaining a term while providing a definition, a lay equivalent, an analogy, or an observable consequence can instead support both immediate comprehension and longer-term medical literacy. It is therefore necessary to distinguish lexical simplification from conceptual accessibility – a distinction that becomes particularly important when expert knowledge is recontextualised for patients or mixed audiences – and to assess accessibility in relation to the explanatory design as a whole.

2.2 Intralingual Captions and Multimodal Accessibility

Intralingual subtitles and captions clearly create a materially observable relationship between spoken discourse and written verbal representation. Subtitling research has demonstrated that captions are shaped by temporal duration, reading speed, segmentation and synchronisation (Díaz Cintas & Remael, 2007, 2020), either condensing or reformulating speech. Although professionally prepared captions may also closely reproduce speech, as in the cases analysed here, such correspondence does not imply an absence of mediation or mere redundancy: subtitles and captions provide an additional route to verbal information and support viewers in adverse listening conditions or with heterogeneous linguistic competence.

The presence of subtitles or captions, however, does not mean that they are the principal source of simplification. If the voice-over already alternates between technical and everyday terms, the written text may preserve that mediation rather than create it. Similarly, animation may supply the spatial or dynamic information required to understand an utterance. Analysis must therefore examine whether information is repeated across modes, divided between them, or made interpretable only through their combination.

In this article, ‘distributed multimodal explicitation’ describes cases in which explanatory meaning is allocated across verbal and visual resources. No single channel contains the complete explanation: for instance, a medical term may appear on screen, a lay gloss may be spoken, and a moving representation may demonstrate the relevant mechanism. This construct is compatible with a broad understanding of accessibility as the removal of communicative barriers, whether these arise from disability, limited domain-specific knowledge, or limited linguistic competence (Greco, 2016, 2018; Rizzo, 2019; Canepari, 2023). It also resonates with Díaz Cintas and Remael’s (2007, pp. 13–14) association between translation and accessibility, since both practices can provide wider access to informational and communicative events.

2.3 Multimodal Mediation and Analytical Dimensions of Accessibility

Multimodal discourse analysis treats meaning as the outcome of relations among semiotic modes rather than as a property of language alone (Jewitt et al., 2016; Kress & van Leeuwen, 2020). Modes have different affordances: speech can sequence causal relations, written labels can stabilise terminology, visual composition can organise comparison and salience, animation can render movement and spatial relations observable, and gestures can direct attention or establish interpersonal involvement. Their contribution must therefore be assessed both individually and through their temporal and compositional coordination.

Four related dimensions of accessibility guide the present analysis. Cognitive accessibility concerns the intelligibility of concepts and relations, as supported by definition, reformulation, analogy, causal sequencing, and visualisation. Sensory accessibility concerns the availability of information through more than one perceptual channel; in the present materials this access remains partial, since captions make speech visually available but do not guarantee full equivalence across modes. Affective accessibility concerns recognition, emotional validation, relatability, voice ownership, and interpersonal positioning. Finally,

operational accessibility refers to the conversion of information into possibilities for action, including symptom recognition, self-management, and guidance on seeking professional assistance.

These dimensions are understood here as analytical descriptors of ‘projected accessibility’ rather than as measures of audience response: they identify textual and audiovisual features that may facilitate access, but do not in themselves demonstrate effects on comprehension, confidence, or engagement. Such effects require reception-based validation.

3. Materials and Method

3.1 Case Selection and Data

The corpus was compiled according to a purposive sampling strategy designed to ensure methodological consistency and comparability, while encompassing distinct forms of medical communication that differ in their communicative purposes, target audiences, and degrees of specialisation.

Several criteria guided corpus compilation: the nature and communicative function of each source; the provenance and quality of subtitles, captions, and transcripts; thematic comparability; and the availability and suitability of the materials for textual processing. The four resources were selected from British healthcare or academic contexts, use British English, provide English-language textual support, are professionally produced, and address respiratory health and disease. Their systematic differences in institutional setting, duration, scope, intended audience, and modal design make it possible to compare communicative strategies without implying statistical representativeness.

Candidate resources were identified through institutional websites and academic audiovisual platforms, and their principal bibliographic and audiovisual metadata – including title, producer, date, duration, source, and URL – were recorded during screening. Selection was conducted manually according to the criteria outlined above. Automatically generated captions were excluded. Where producer- or platform-supplied subtitles, captions and transcripts required conversion or preprocessing, the resulting working texts were checked manually against the audiovisual source, especially in passages containing medical terminology or discrepancies among speech, subtitles, captions, and written labels. Retrieval and preprocessing tools facilitated data preparation but did not determine source selection, analytical categorisation, or interpretation.

Given the differences in duration and pedagogical scope among the four resources, comparison was not based on equal-duration samples or on the assumption of complete content equivalence. The unit of comparison was the functionally comparable explanatory sequence: a segment devoted to naming, defining, visually representing, or causally explaining respiratory anatomy, physiological processes, symptoms, or disease mechanisms. The two short institutional videos were analysed in full. Qualitative analysis of *How We Breathe* drew on functionally relevant sequences distributed across the complete resource, including later sections on obstruction and anxiety, while its temporal profile was calculated over the complete narrated span. Analysis of the substantially longer HSTalks lecture was

restricted to the coherent opening sequence from 00:00 to 08:29. This asymmetric delimitation reflects the different structures of the resources and is reported explicitly rather than treated as equivalent sampling.

Within the first academic–institutional pair, *How We Breathe* was produced by the International Primary Care Respiratory Group as an educational resource for healthcare professionals to use with patients. The complete video lasts 23 minutes and 36 seconds and is divided into sections covering definitions of breathing, ventilation, the bronchial tree, respiration, breathlessness, obstruction, restriction, deconditioning, anxiety, breathing pattern disorders, airborne particles, and a final summary. The resource explicitly states that it combines medical terminology with lay vocabulary. Medical terms are presented in visual labels, while both specialised and non-specialised registers are used in the voice-over.

The NHS *Asthma* video lasts approximately 2 minutes and 41 seconds. It defines asthma, introduces basic airway anatomy, explains triggers and airway narrowing, and relates these processes to wheezing, mucus production, and coughing. It also grades the severity of asthma and distinguishes preventive self-management from the medical intervention required during an acute attack. Throughout this explanation, terminology is retained selectively and supported by disease-specific animation. Its compact scope therefore makes the video a useful institutional counterpart to the broader pedagogical design of the IPCRG resource.

The data comprise audiovisual files, English voice-over transcripts, curated intralingual subtitles, captions or platform transcripts, slides where available, and time-coded segmentation. *How We Breathe* was examined using a 1920 × 1080 copy, separate audio, selectable English (United Kingdom) captions, and a detailed time-coded transcript. The NHS *Asthma* resource was examined in its entirety using the complete video, the curated subtitles, and the transcript embedded in the institutional webpage. Distinguishing these forms of textual evidence reduces the risk of treating speech-recognition errors, faulty segmentation, or terminological inaccuracies as deliberate editorial or translational choices.

For the COPD pair, the academic material consists of the opening 8 minutes and 29 seconds of Barnes’s HSTalks lecture, together with its curated subtitles, transcript and downloadable slides. This extract constitutes a coherent macrosequence, progressing from the clinical differences and shared features of asthma and COPD to their pathology, the mechanisms of airway obstruction, and an introduction to their cellular mechanisms. It concludes with a consolidated comparison of the inflammatory patterns associated with the two conditions.

The institutional material comprises the complete 2-minute-41-second Blackpool Teaching Hospitals video *Symptoms of COPD*. It was examined through the audiovisual file, its on-screen verbal material, and a manually prepared working transcript checked against the recording. The academic extract employs a lecturer-led slide format, whereas the institutional video combines a visible healthcare presenter, emphatic hand gestures, concise written headings, and illustrative inserts.

Table 1. Overview of sources, analytical spans, and textual evidence

Resource	Context and audience	Total duration	Qualitative span	Textual evidence	
IPCRG How We Breathe	Academic; professional–patient education	23:36	Selected sequences throughout	Selectable (UK) time-coded transcript	English captions;
NHS Asthma	Institutional; patients and public	c. 2:41	Complete video	Curated transcript embedded in webpage	subtitles;
HSTalks Barnes	Academic; specialist audience	44:00	00:00–08:29	Curated platform slides	subtitles and transcript;
Blackpool Symptoms of COPD	Institutional; patients and public	2:40.8	Complete video	On-screen manually prepared and verified transcript	text; working

3.2 Analytical Procedure

For the purposes of this study, the analysis combines linguistic, subtitle- and caption-based, temporal, and visual annotation. First, spoken and captioned verbal material was segmented according to explanatory function. Medical terms, lay equivalents, definitions, analogies, causal links, epistemic markers, and self-management formulations were identified. Second, speech–caption relations were examined in terms of verbatim reproduction, reduction, and reformulation. Third, visual resources were annotated for written labels, anatomical anchoring, motion, spatial perspective, slide organisation, causal visualisation, and coordination with verbal deixis; where a presenter was visible, emphatic gesture and audience positioning were also considered. Finally, findings were compared first within each pair and then across pairs in relation to scope, projected audience, and communicative purpose.

The main analytical categories taken into account were terminological retention, parallel lexical labelling, local explicitation, conceptual differentiation, analogy, causal chaining, embodied reformulation, visual anchoring, dynamic causal visualisation, subtitle and caption synchronisation, affective positioning, and self-management orientation. These features were subsequently related to the four accessibility dimensions defined in Section 2.3. Speech rate was treated as one dimension of temporal accessibility rather than as an isolated index of comprehensibility, since a moderate rate does not necessarily compensate for unexplained conceptual density, whereas faster delivery may remain manageable when segmentation and visual support are strong.

Temporal analysis was based on time-coded English transcripts and verified against the

audiovisual recordings. Word counts included lexical tokens in spoken narration and excluded section headings, credits, and non-verbal written labels. Global speech rate was calculated as words per minute over elapsed narrated time and therefore includes pauses; it is distinguished here from articulation rate, which would be calculated over speaking time alone. Selected authored sections of *How We Breathe* were also measured separately to examine internal variation. Silent intervals of at least 0.4 seconds were identified with FFmpeg's *silencedetect* filter and checked against transcript boundaries and audiovisual transitions. Non-silent time was calculated as elapsed duration minus detected silence and expressed as a proportion of the analysed span. This acoustic measure may include non-speech sounds and should not be interpreted as a direct measure of speaking time. Because file provenance, soundtrack level, and recording conditions differ, pause measures are interpreted descriptively and only within the relevant pair. Quantitative results are consequently presented in pair-specific tables rather than as a single cross-source ranking.

The framework thus distinguishes four interacting levels: verbal selection, caption realisation, visual explanation and accessibility function. This prevents an isolated feature from being interpreted automatically as evidence of accessibility. For example, a technical term is evaluated together with the definition, label, animation or causal context that supports it.

Figure 1 summarises the relation among the paired sources, communicative modes, analytical dimensions, comparative parameters, and accessibility functions.

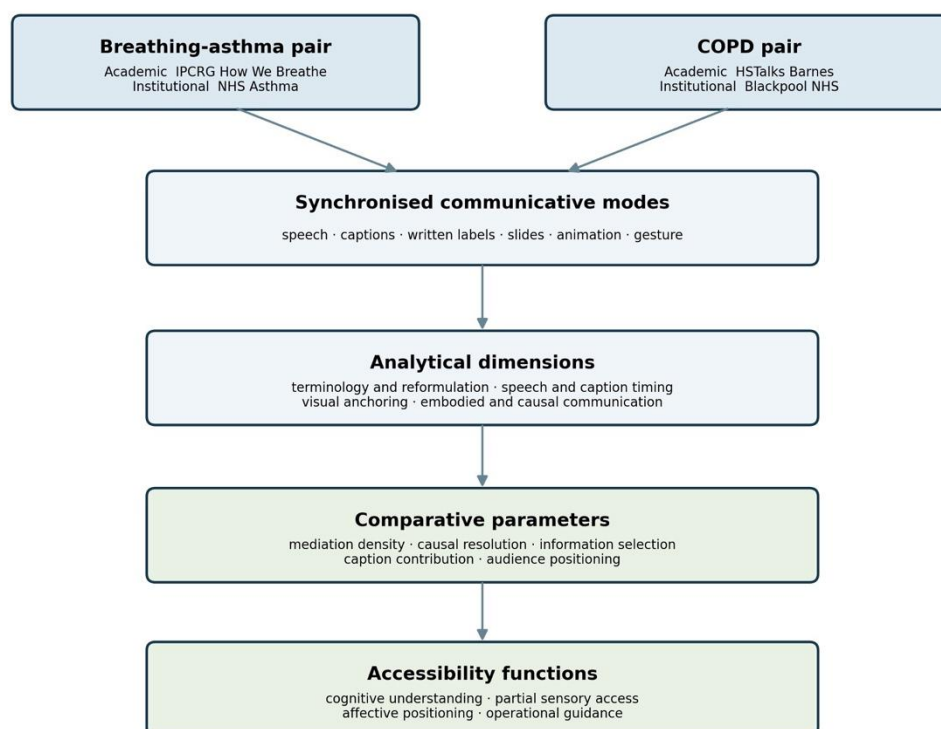


Figure 1. Comparative multimodal analytical framework for the four-source respiratory case study

4. Findings

4.1 Dual Register and Selective Terminological Retention

The most distinctive feature of *How We Breathe* is its explicit dual-register design. The introduction announces the use of both medical terminology and lay vocabulary and explains the role of animation in the video (00:00:05–00:01:04). This dual-register principle is realised through recurrent pairings such as “thorax/chest” (00:03:02), both of which occur in the speech, captions, and on-screen labels. By contrast, the pair “abdomen/belly” is spoken and captioned but does not appear as an on-screen label (00:03:05), while “intercostal” is preceded in speech by “between our ribs” (00:03:13). Similarly, “sternum” is glossed as “breastbone” (00:03:24), and “pharynx” is translated into the plain-English equivalent “throat” (00:06:52–00:06:54). “Larynx” is immediately followed by the more familiar “voice box” (00:06:57–00:06:58), while “trachea” is reformulated as “windpipe” (00:07:14–00:07:15). In these instances, the selected terms do not function as incidental synonyms. Rather, each pairing enables viewers to associate a specialised label with a familiar conceptual referent.

The ordering of each pair varies, but the underlying operation remains stable: terminological precision is retained while immediate access is supplied locally. This strategy, described here as ‘parallel lexical labelling’, also occurs when on-screen text uses acronyms or chemical symbols. For instance, the narration first identifies the extended term “carbon dioxide” (00:02:03), after which its chemical formula “CO₂” appears on screen and is introduced as “written as” (00:02:07). Unlike replacement-based simplification, this strategy preserves both registers, making the specialised term and its more familiar equivalent simultaneously available to the viewer. Accessibility is therefore constructed not by removing medical terminology, but by establishing an explicit correspondence between expert and lay forms. From a Medical English perspective, such dual-register mediation may support immediate comprehension while also promoting terminological learning: viewers can understand the concept through the familiar expression and, at the same time, associate it with the term they may subsequently encounter in clinical consultations, medical documents, or other health-information resources. The strategy is thus consistent with the video’s instructional purpose, which is not merely to reduce conceptual difficulty but to help patients engage with medically accurate vocabulary.

More complex physiological processes, such as “expiration or exhalation” (00:04:25), are mediated through speech, captions, and animation. Expiration is defined as the phase of ventilation in which air leaves the lungs and the respiratory muscles relax, allowing the lungs and airways to recoil after being stretched open (00:04:27–00:04:36). Similarly, the specialised term “tidal breathing” (00:05:25) is introduced only after the process has been described in concrete, quantitative terms: “When we’re relaxed, adults breathe in and out approximately eight to 18 times per minute, and each breath contains about 500 millilitres of air” (00:05:17–00:05:25).

On other occasions, spoken and written language retain specialised terminology without verbal scaffolding, as with “scalene muscles” (00:03:20) and “pleural membranes” (00:03:39). Here, however, selective colour and animation identify the anatomical structures and make

their spatial relations more interpretable. Pastel-coloured structures stand out against a darker background, directing attention to the body and to the parts currently under discussion.

How We Breathe also differentiates among “breathing”, “ventilation”, and “respiration” (00:01:23–00:02:16). Ventilation is defined as the muscular work that moves air into and out of the lungs (00:01:48–00:01:50); breathing as the process that brings oxygen into the body and removes carbon dioxide (00:01:54–00:02:00); and respiration as the process through which oxygen is supplied to cells and tissues (00:02:13–00:02:16). Thus, although this distinction increases the number of technical concepts introduced, it reduces ambiguity and supports a more precise mechanistic understanding.

The sequencing is also important: in the asthma-related academic sequence, technical naming often follows accessible conceptual preparation. “Muscle tightening” (00:12:45) is introduced before “bronchospasm”, while the description of the muscles tightening around the bronchi and bronchioles provides a conceptual basis for the related term “bronchoconstriction” (00:12:51). The sequence can therefore be represented as a progression from a plain-language mechanism to a clinical example, a specialist label, and further anatomical elaboration, and through this progression, terminology becomes the endpoint of the explanation rather than its prerequisite.

The video also recruits bodily experience: viewers are invited to inhale, hold their breath, and attempt another inhalation to approximate the sensation of air trapping. This invitation to ‘learn by doing’ recalls experiential-learning traditions (Dewey, 1938; Kolb, 1984), while the colloquial formulation “breathing is just a slog” (00:14:40) connects increased work of breathing with everyday experience. Animation, instruction, and the viewer’s body may consequently operate together to support understanding.

The NHS video uses comparable resources more selectively. It likewise has no visible speaker: a voice-over accompanies anatomical images and animated physiological processes against a dark background. Verbally, however, it selects “windpipe” (00:00:38) without adding ‘trachea’, while retaining “bronchi”. It also introduces the specialist adjective “atopic” together with the explanation that affected people have an “allergic-type reaction to external triggers” (00:01:08–00:01:15). “Wheezing” (00:01:50), which *How We Breathe* introduces through the adjective “wheezy” without a local definition (00:10:49), is rendered perceptually meaningful as a “whistling sound” (00:01:49–00:01:50). On this occasion as well, captions closely reproduce the voice-over, while animation clarifies the mechanisms described, thereby mediating medical concepts and making them interpretable.

Unlike *How We Breathe*, the NHS video does not preserve the full distinction among breathing, ventilation, and respiration, treating the first and third as locally interchangeable in the phrase “during breathing, or respiration” (00:00:27–00:00:29). The distinction between pulmonary ventilation and cellular respiration is not required by the video’s immediate explanatory purpose. Indeed, this strategy, described here as ‘functional terminological reduction’, retains the technical term ‘respiration’ while omitting a physiologically relevant contrast that is not essential to the communicative task at hand.

Furthermore, although the NHS video is less interactive, it uses sensory explication when wheezing is described as a whistling sound. This converts a clinical symptom into something potentially recognisable. Both resources therefore mediate terminology through experience, although one elicits embodied simulation and the other supplies a perceptual gloss. Institutional mediation thus removes some distinctions while retaining clinically useful terms whose conceptual burden is controlled through local explanation and visual representation.

The contrast between the two sources cannot therefore be reduced to technical language vs plain language: rather, it could be described in terms of systematic dual naming vs economical selective retention, where the former constructs a broad lexical repertoire, whereas the latter selects the minimum terminology required for a disease-specific mental model.

This comparison therefore illustrates why reduced complexity should not automatically be equated with lexical simplification. Compared with *How We Breathe*, the NHS video seeks to support accessibility primarily through scope restriction and a lower degree of causal and anatomical resolution. Although it retains specialist items such as “bronchi” (00:00:42), “atopic” (00:01:10), “inflammation” (00:02:20), and “acute asthma attack” (00:02:31–00:02:32), it does not address the more detailed physiological and anatomical distinctions presented in the academic resource, including the role of the “pleural membranes” (00:03:39).

In the COPD pair, the HSTalks extract constructs disciplinary authority through a stable lecture format in which Barnes’s spoken explanation is synchronised with dense but highly organised slides. The resource constructs accessibility primarily through contrastive architecture. Asthma and COPD are repeatedly placed in parallel columns; colour highlighting guides attention; histological images anchor pathological differences; and cellular diagrams spatialise relations among allergens, irritants, inflammatory cells and physiological effects. The curated captions closely follow specialist speech and therefore provide linguistic reinforcement rather than extensive intralingual reformulation. Thus, the configuration supports conceptual discrimination but presupposes substantial biomedical literacy.

Specialist terminology is retained throughout, but it is frequently embedded in systematic comparison. For example, “variable and intermittent wheezing” (00:00:20–00:00:26) in asthma is contrasted with “continual symptoms and in particular shortness of breath on exertion” (00:00:29–00:00:34) in COPD; reversibility and responsiveness to bronchodilators and steroids are contrasted with largely fixed obstruction and poorer treatment response (00:01:39–00:01:55). From 00:03:02, the pathology slide contrasts “eosinophils” (00:03:34) in asthma with “neutrophils” (00:03:41) in COPD. It also contrasts smooth-muscle cell hyperplasia (00:03:50–00:03:52) and hypertrophy (00:03:58) in asthma with “minimal thickening of airway smooth muscle” (00:04:04–00:04:08) in COPD. Similarly, the speaker compares the thickening of “the basement membrane under the airway epithelium” (00:04:11–00:04:15) found in asthma with “peribronchiolar fibrosis” (00:04:41) and emphysematous “alveolar disruption” (00:04:47–00:04:49). The terms are not simplified, although their interpretation is scaffolded by repetition, parallel syntax, spatial alignment and

visual contrast.

The Blackpool video distributes information differently. A visible healthcare presenter accompanies her speech with recurrent hand gestures that emphasise key points, increase their visual salience, and may help sustain viewer engagement. Brief written headings and illustrative inserts further isolate recognisable symptoms and risk-related information. Typographically separated from the surrounding material and displayed on individual slides against a white background, the headings foreground the topic being addressed and direct viewers' attention.

The sequence moves from “shortness of breath” (00:00:22) and “cough” (00:00:49), through “irritation of the airways” (00:00:59) and the contrast between healthy and damaged cilia (00:01:04–00:01:11), to other symptoms such as “wheeze” (00:01:21), which is glossed both in speech and in the on-screen text as an “abnormal whistling sound” (00:01:28); “tiredness” (00:01:31); and “loss of appetite leading to weight loss” and “ankle swelling” (00:01:54–00:01:58). The video subsequently introduces “triggers” (00:02:00), including “smoking, air pollution or dust, cold or very hot weather and chest infections” (00:02:14–00:02:16).

The closing recommendations – “Stop smoking” (00:02:22) and “For advice, please speak to your practice nurse or doctor” (00:02:27–00:02:33) – convert symptom recognition into an explicit pathway for action, shifting the video's predominant communicative function from informative to conative. Rather than constructing an integrated pathological model, the video therefore proceeds through small, clearly delimited units of information. Cognitive accessibility and actionability are supported by segmentation, concrete images, and concise headings, while the presenter's emphatic gestures may contribute to affective and interpersonal accessibility by directing attention, reinforcing verbal emphasis, and potentially establishing a more immediate relationship with the viewer.

4.2 Analogy and Causal Resolution

Analogy may itself be understood as a form of accessibility because it maps unfamiliar or specialised phenomena onto experiential domains that audiences already know how to read, see, or imagine. In line with Lakoff and Johnson's (1980) account of metaphor in everyday conceptualisation, analogical mediation is not merely ornamental: it provides a cognitive structure through which viewers can approach microscopic, abstract, or otherwise imperceptible medical processes. It organises perception by connecting unfamiliar phenomena with familiar spatial, bodily, and causal patterns. Its accessibility value is therefore closely linked to relatability, understood here as the possibility of connecting specialist knowledge with recognisable actions, sensations, objects, or situations. Analogy does not necessarily simplify the underlying concept; rather, it creates an experiential route into it through the coordinated work of language, image, movement, and sound. The term is used here in a broad analytical sense that also encompasses metaphorical mappings.

In audiovisual health communication, such mappings may be activated verbally, visually, or through the interaction of modes, as when airflow obstruction is represented as a blocked passage or muscular deconditioning as a downward spiral. The accessibility potential of such

mappings lies not only in verbal replacement but also in enabling viewers to connect expert knowledge with familiar perceptual and experiential frames.

How We Breathe, for instance, uses analogical and metaphorical mappings to relate unfamiliar physiological processes to everyday experiential schemas. Ventilation is compared to ventilating a house; the bronchial tree is imagined as an upside-down oak tree; the brainstem becomes a signal box; and mitochondria are described as cellular energy factories. Specialist terminology is therefore retained, while analogy builds conceptual bridges between the medical concepts and familiar domains of experience.

By contrast, the NHS resource relies less on analogy and more on a short causal chain: an asthma trigger produces airway-wall swelling and muscular contraction; the airway narrows; airflow creates wheezing; mucus further obstructs the lumen and causes coughing; and breathing becomes more difficult. The narration and animation transform symptoms from a list into consequences of visible physical changes.

The two resources therefore differ in causal resolution: the NHS selects a short, clinically sufficient sequence, whereas IPCRG extends the sequence from early airway closure to trapped stale air, hyperinflation, mechanically shortened respiratory muscles, increased muscular effort, and worsening breathlessness. This difference reflects communicative purpose: institutional communication provides a usable disease model, whereas respiratory education develops a more transferable physiological model.

The COPD pair relocates causal explanation rather than merely shortening it. In HSTalks, the causes of airway obstruction are differentiated mechanistically: asthma is associated principally with bronchoconstriction caused by smooth-muscle contraction, with oedema and mucus plugging as additional mechanisms (00:05:13–00:05:43); COPD is associated with small-airway fibrosis and the loss of alveolar attachments in emphysema, which allows airways to close on expiration (00:05:56–00:06:11). The subsequent cellular-mechanism sequence (00:06:32–00:08:29) extends the chain from allergens or irritants to different cells and physiological effects.

By contrast, Blackpool does not reproduce this pathological chain. Instead, it reorganises causality around observable consequences, environmental triggers and actionable advice. The contrast is therefore between expert causal differentiation and symptom-oriented causal compression, not between the presence and absence of explanation.

4.3 Distribution Across Voice, Captions and Visual Representation

A clear instance of ‘distributed multimodal explication’ occurs in the introduction to *How We Breathe*, which makes its semiotic division of labour explicit. Visual labels favour medical terminology, the voice-over moves between medical and lay vocabulary, and moving images provide spatial access to bodily structures. The design is also metacommunicative, telling viewers both what they will learn and how the different modes are organised to guide interpretation.

Several sequences employ deictic prompts, including “notice how high up the lungs reach”

(00:05:55–00:05:58) and “notice the rib cage goes right round the body” (00:06:07–00:06:11). These prompts direct visual attention and encourage viewers to interpret the anatomy as a three-dimensional structure. Animation therefore does more than illustrate an independently complete verbal description: it supplies depth, movement, and spatial relations that speech alone does not fully encode.

Dynamic processes are especially dependent on this coordination. Diaphragm movement, ribcage expansion, lung recoil, airway narrowing, air trapping and particle deposition unfold over time. Spoken causal syntax structures the sequence, captions stabilise access to the terminology and animation makes the mechanism observable. This configuration exemplifies distributed multimodal explication.

The NHS video applies the same principle in a more compressed form. Its visual progression follows air through the upper airways and bronchi (00:00:34–00:00:44) before depicting airway swelling, muscular contraction, narrowing, and mucus accumulation (00:01:36–00:01:58). The voice-over names each process, while the corresponding stages are presented on screen as headings and the animated airway renders the development of obstruction visible. Here, ‘distributed multimodal explication’ depends primarily on temporal alignment: viewers hear the causal sequence while simultaneously watching the corresponding anatomical changes unfold.

Captioning must consequently be interpreted as one component of the wider communicative design. In the IPCRG resource, the curated English track largely preserves a voice-over that already contains intralingual mediation. The captions provide textual access to the spoken content and reinforce medical terminology, but the principal shift between expert and lay discourse frequently occurs within the spoken text, prior to captioning. This finding therefore cautions against treating every intralingual caption as an autonomous translation whose accessibility depends chiefly on lexical change.

The COPD pair reveals two further configurations of explanatory labour. In the HSTalks lecture, the lecturer’s speech provides causal and evaluative continuity, the captions stabilise dense terminology, and the slides externalise comparison through columns, colour, and images. In the Blackpool resource, by contrast, the visible presenter coordinates speech and emphatic gestures, while concise headings and illustrative sequences organise the verbal message into symptom- and action-centred units. Consequently, in neither case can accessibility be assigned to one channel in isolation.

4.4 Speech Rate and Temporal Organisation

In relation to the first pair, the time-coded transcripts establish a marked difference in delivery rate. *How We Breathe* contains approximately 3,136 spoken words across a narrated span of 22 minutes and 33 seconds, producing a global rate of 139.1 words per minute (WPM). The NHS Asthma transcript contains approximately 260 words across 2 minutes and 40 seconds, producing a rate of approximately 97–98 WPM. These values measure total elapsed time and therefore include pauses; as such, they should be described as speech rate rather than articulation rate.

Table 2 reports temporal measures for the breathing–asthma pair only. This placement is deliberate: the measures support the within-pair interpretation and are not treated as directly commensurable with the differently delimited COPD materials.

Table 2. Speech-rate measures for the breathing–asthma pair

Source or selected section	Words	Duration	Speech rate
IPCRG complete narrated span	3,136	22:33	139.1 WPM
Introduction	166	1:23	120.0 WPM
Definitions	122	0:56	130.7 WPM
Ventilation	505	3:47	133.5 WPM
Bronchial tree	286	2:04	138.4 WPM
Respiration	218	1:26	152.1 WPM
Obstruction	656	4:29	146.3 WPM
Stress and anxiety	215	1:24	153.6 WPM
Particles	295	1:48	163.9 WPM
Summary	79	0:45	105.3 WPM
NHS Asthma	260	2:39–2:41	97–98 WPM

Silence detection using the procedure described in Section 3.2 identified 315 silent intervals within the IPCRG narrated span at the 0.4-second threshold (median 0.74 seconds; 109 intervals of at least one second) and 49 in the NHS clip (median 0.85 seconds; 22 intervals of at least one second). Because silence detection is sensitive to soundtrack level and recording conditions, these values support qualitative interpretation rather than direct acoustic ranking. The stable result is that the NHS video combines a substantially lower global speech rate with frequent segmentation, whereas IPCRG varies tempo according to explanatory function and slows markedly in its summary.

Temporal organisation therefore reinforces the contrast between elaborative and compressed mediation. The IPCRG narration averages 139.1 WPM but varies substantially by section, from 105.3 WPM in the concluding summary to 163.9 WPM in the particle sequence. The asthma-relevant obstruction section is delivered at 146.3 WPM. By contrast, the NHS narration averages approximately 97–98 WPM. Its limited conceptual scope and slower delivery may reduce processing pressure. IPCRG combines a faster and more variable tempo with paired terminology, repetition, visual labels, analogy, and animation. Speech rate

consequently functions as one component of mediation density rather than as an autonomous accessibility measure.

For the COPD pair, the 00:00–08:29 HSTalks extract contains approximately 941 words after normalising obvious spacing artefacts in the curated transcript. Across the elapsed span of 509 seconds, this produces a speech rate of 110.9 WPM. At a 0.4-second acoustic threshold, the extract contains 92 silent intervals, totalling 61.4 seconds; their median duration is 0.53 seconds, and six intervals last at least one second. The corresponding non-silent-time proportion is approximately 87.9%. These results describe a relatively continuous lecture delivery whose pauses chiefly accompany transitions within and between contrastive slides.

The complete Blackpool video lasts approximately 160.8 seconds. Using the same acoustic procedure, it contains 34 silent intervals totalling 53.7 seconds; their median duration is 0.60 seconds, and seven intervals last at least one second. The non-silent-time proportion is therefore approximately 66.6%. Although the institutional video contains fewer pauses in absolute terms, silence occupies a substantially larger share of its much shorter duration. This pattern corresponds to its segmented organisation: headings, illustrative inserts, and changes of informational unit interrupt the presenter-led delivery and may provide additional processing opportunities.

The available Blackpool transcript contains 334 words. Calculated over the complete 160.8-second recording, this gives a global speech rate of approximately 124.6 WPM, rounded to 125 WPM. Pause values remain descriptive because recording levels and the presentational use of silence can affect acoustic detection.

Table 3. Temporal measures for the COPD pair

Source	Analysed span	Words	Speech rate	Pauses ≥ 0.4 s	Total silence	Median pause	Non-silent time
HSTalks	8:29	c. 941	110.9 WPM	92	61.4 s	0.53 s	87.9%
Blackpool	2:40.8	334	125 WPM	34	53.7 s	0.60 s	66.6%

The comparison confirms that speech rate cannot be interpreted independently of pausing and temporal segmentation. Although the Blackpool video has a higher global speech rate than the HSTalks lecture (approximately 125 WPM versus 110.9 WPM), detected non-silent time accounts for only 66.6% of its total duration, compared with 87.9% in the academic resource. The audiovisual analysis indicates that its delivery is organised into relatively rapid stretches of speech separated by pauses and presentational transitions, whereas the HSTalks lecture maintains a slower but more continuous verbal flow. In the Blackpool video, these pauses, together with concise headings, concrete images, and gestural emphasis, may create additional opportunities for processing despite the faster delivery of individual spoken segments. Temporal accessibility thus depends not on speech rate alone but on its interaction

with pausing, segmentation, and multimodal support.

4.5 Cognitive, Sensory, Affective, and Operational Accessibility

Both resources in the first pair deploy verbal and visual strategies that may increase linguistic and cognitive accessibility. At the sensory level, curated captions make spoken information visually available, while voice-over, written labels, and animation distribute explanatory content across different perceptual channels. This cross-modal distribution provides partial sensory accessibility, although it does not imply that all information is fully accessible through each mode independently.

How We Breathe also incorporates the emotional experience of breathlessness into its physiological account. In a later sequence selected for its relevance to affective accessibility, it explains how breathlessness may produce fear, panic, and thoughts of dying and how these responses may increase muscular tension and alter breathing, thereby further intensifying symptoms (00:18:42–00:18:56). Naming this feedback loop acknowledges distress while making it cognitively interpretable and relatable. The subsequent recommendation of breathing exercises translates this explanation into a possible response (00:18:56–00:19:02). Thus, although the video contains neither a talking head nor first-person testimony, its discourse may support affective accessibility while linking emotional recognition to operational guidance.

The NHS resource is more emotionally restrained. Its opening formulation states that asthma is chronic and cannot be cured but can be managed, placing agency immediately after incurability through an adversative structure. Later, severity is graded from mild to life-threatening without implying that deterioration is inevitable. Preventive treatment is associated with patient self-management, whereas an acute attack is linked to medical intervention according to its severity.

Both resources therefore connect knowledge with agency, but they construct this pathway differently. The IPCRG resource is designed to support confidence, communication, and appropriate breathing techniques through understanding. The NHS video compresses the pathway into symptom recognition, preventive management, and escalation when necessary. Accessibility is therefore constructed across cognitive, affective, sensory, and operational dimensions, although these receive different degrees of emphasis in each resource.

The COPD comparison also demonstrates why multimodality should not automatically be equated with affective accessibility. In the absence of speaking roles for patients or relatives, the two resources construct accessibility through different forms of explanation and audience positioning. The HSTalks lecture primarily supports cognitive accessibility through epistemic authority, systematic explanation, and schematic visual comparison. The Blackpool video may instead enhance relatability through recognisable symptoms, direct address, facial visibility, and embodied gestural emphasis. These features may reduce interpersonal distance, but they do not provide the identification, validation, or emotional reassurance that can be generated by lived-experience testimony. Affective accessibility should therefore be treated as a distinct analytical dimension, supported by voice ownership, narrative positioning, and

opportunities for recognition rather than inferred from human visibility alone.

5. Comparative Discussion

The comparison emphasises that, across the four sources, academic and institutional videos do not occupy opposite ends of a simple technical-to-plain *continuum*. All retain some specialist terminology, and all rely on interactions between verbal and visual modes, but they organise access differently. Their central differences lie in how much knowledge they select, how many explanatory steps they preserve, and how explicitly they teach viewers to interpret multimodal representations. *How We Breathe* constructs an extensive, layered model; NHS *Asthma* compresses its causal account around a single condition; HSTalks adopts a contrastive disciplinary architecture; and Blackpool segments information around symptoms, triggers, and action.

Two qualitative analytical parameters emerge from these cases. The first is ‘mediation density’, understood as the number and variety of explanatory supports deployed around a specialised concept rather than as a validated numerical measure. *How We Breathe* combines high information density with high mediation density through definitions, paired labels, analogies, captions, guided viewing, and bodily simulation. NHS *Asthma* has lower information density and employs fewer but tightly coordinated supports. HSTalks combines high terminological density with repeated parallelism, colour coding, and visual comparison, whereas Blackpool maintains lower verbal and conceptual density while supporting interpretation through segmentation, concrete imagery, and embodied presentation. Complexity therefore cannot be inferred from word count or terminology alone.

The second parameter is ‘causal resolution’, understood qualitatively as the number of intermediate causal relations retained in the explanation of a mechanism. NHS *Asthma* selects a short causal chain sufficient to support an understanding of symptoms and management, whereas *How We Breathe* preserves longer physiological sequences. HSTalks differentiates the cellular and pathological mechanisms of asthma and COPD, while Blackpool privileges recognisable consequences and practical action. Causal compression can therefore constitute a form of intralingual mediation even when no individual word is simplified.

These findings refine the relationship between intralingual translation and captioning. A speech–caption comparison may reveal little lexical transformation even when the audiovisual text as a whole contains extensive intralingual mediation. As the preceding analyses show, reformulation may already be incorporated into the voice-over through parallel naming, definition, analogy, and experiential language, or into the organisation of slides and headings. Captions then transmit and stabilise this already mediated discourse. The location of translation must therefore be established empirically rather than assumed from the existence of a written target channel.

Furthermore, the visual mode introduces a related complication, since animation may explicitate a spatial or causal relation without adding words. For instance, when airway narrowing is shown while wheezing is explained, the visual sequence translates an invisible physiological mechanism into an observable event. In this way, intersemiotic mediation

changes the interpretability of the accompanying intralingual verbal material, demonstrating that linguistic and intersemiotic strategies are analytically distinguishable yet functionally interdependent.

Finally, the temporal findings also complicate any equation between accessibility and linguistic simplicity. NHS *Asthma* combines a slower speech rate with causal compression, whereas *How We Breathe* presents more information at a faster rate while creating multiple routes into meaning. Its accessibility design can therefore be described as ‘compensatory and layered’ rather than primarily reductive. The COPD comparison points in the same direction: HSTalks uses visual parallelism to organise dense specialist exposition, while Blackpool uses short sequential units, presentational pauses, and continuous human presence to structure processing. Potential processing pressure should therefore be evaluated as a relation among information density, mediation density, temporal segmentation, and modal coordination.

6. Limitations and Future Research

This study does not claim that any of the resources produces measurable gains in comprehension. Accessibility is inferred from the organisation of potentially facilitative resources rather than tested through reception. Nor does the comparison rank the videos against a universal standard. The academic resources target clinicians, trainees, or mixed audiences seeking mechanistic literacy, whereas the institutional resources target viewers seeking immediate explanation, symptom recognition, or practical guidance. Accessibility therefore remains dependent on communicative purpose and intended audience.

The small purposive sample constitutes a further limitation. Four cases cannot represent academic and institutional medical discourse in general, while asymmetries in duration, condition coverage, and modal design preclude straightforward quantitative comparison. These asymmetries are nevertheless analytically meaningful, since scope, duration, and information selection are themselves components of communicative design.

The temporal analysis was limited to speech rate and pausing and did not include a systematic examination of prosody. Intonation, stress, pitch range, and rhythm may nevertheless contribute to accessibility by segmenting information, foregrounding key terms, communicating stance or reassurance, and sustaining interpersonal engagement. Their analysis, however, would require a dedicated acoustic and perceptual protocol and therefore lies beyond the scope of the present study.

Accordingly, the accessibility discussed in the present study should be understood as ‘projected accessibility’, that is, accessibility expected on the basis of the linguistic, temporal, and multimodal affordances identified in the materials, rather than as an empirically demonstrated audience effect. Given the analytical scope of the study and the space available, no reception-based data were collected. A follow-up study planned for the 2026/2027 academic year will test selected materials with undergraduate students in Foreign Languages, Master’s students in specialised translation, and medical students and trainees, in order to examine how the patterns identified here relate to actual comprehension, recall, perceived accessibility, and engagement. This reception phase will make it possible to assess the

descriptive parameters proposed in the present study against users' responses.

7. Conclusion

This study has examined how four English-language respiratory videos from British healthcare and academic contexts coordinate spoken language, subtitles, captions or transcripts, and visual resources to mediate specialised medical knowledge. *How We Breathe* constructs broad breathing literacy through dual-register design, analogy, embodied explanation, and exploratory animation. NHS *Asthma* constructs a concise disease model through selective terminology, local explication, causal compression, and self-management discourse. HSTalks organises specialist COPD knowledge through systematic contrast, slides, and curated captions, whereas Blackpool uses a visible presenter, emphatic gesture, headings, and concrete inserts to support symptom recognition and action.

The first research question is therefore addressed by the finding that all four resources distribute meaning across modes rather than locating accessibility in simplified wording alone. Technical labels preserve disciplinary precision; spoken reformulation creates lexical and conceptual bridges; captions stabilise access; and animation, slides, images and gestures make spatial, causal or sensory relations available. In relation to the second question, the findings show how communicative purpose affects the scope of selected knowledge, the resolution of causal explanation, temporal organisation and the kinds of accessibility prioritised.

The comparison thus supports a context-sensitive account of audiovisual medical accessibility. Terminological retention is compatible with accessibility when concepts are surrounded by sufficient mediation, while information reduction may occur through selection, segmentation and causal compression rather than lexical substitution. Mediation density and causal resolution provide useful categories for describing these differences, but sensory, affective and operational accessibility must also be distinguished. Intralingual subtitles and captions should therefore be analysed within the multimodal architecture that gives them meaning: they may reformulate speech, but they may also provide access to a discourse whose mediation has already been designed across voice, text, image, embodiment and time.

References

- Barnes, P. (2021, May 30). COPD and asthma: Similarities and differences [Video]. *The Biomedical & Life Sciences Collection, Henry Stewart Talks*. <https://doi.org/10.69645/SGBK6016>
- Blackpool Teaching Hospitals NHS Foundation Trust. (2026). Symptoms of COPD [Video]. Retrieved from <https://healthvideos.bfwh.nhs.uk/copd?videoId=3062>
- Canepari, M. (2023). *Specialized languages and graphic art*. Peter Lang. <https://doi.org/10.3726/b19709>

- Canepari, M. (2024). On the theory and practice of intralingual translation and its typologies. *Perspectives*, 32(6), 1–22. <https://doi.org/10.1080/0907676X.2024.2375219>
- Canepari, M. (forthcoming). *Intralingual translation in action: theories, texts and strategies*.
- Dewey, J. (1938). *Experience and education*. Macmillan.
- Díaz Cintas, J., & Remael, A. (2007). *Audiovisual translation: Subtitling*. St. Jerome Publishing.
- Díaz Cintas, J., & Remael, A. (2020). *Subtitling: Concepts and practices*. Routledge. <https://doi.org/10.4324/9781315674278>
- Greco, G. M. (2016). On accessibility as a human right, with an application to media accessibility. In A. Matamala, & P. Orero (Eds.), *Researching audio description: New approaches* (pp. 11–34). Palgrave Macmillan. https://doi.org/10.1057/978-1-137-56917-2_2
- Greco, G. M. (2018). The nature of accessibility studies. *Journal of Audiovisual Translation*, 1(1), 205–232. <https://doi.org/10.47476/jat.v1i1.51>
- Hill-Madsen, A. (2015). Lexical strategies in intralingual translation between registers. *HERMES Journal of Language and Communication in Business*, 54, 85–105. <https://doi.org/10.7146/hjlc.v27i54.22949>
- Hill-Madsen, A. (2019). The heterogeneity of intralingual translation. *Meta*, 64(2), 537–560. <https://doi.org/10.7202/1068206ar>
- Hill-Madsen, A. (2022). Transformational strategies in diaphasic translation: Three case studies. *Perspectives*, 30(4), 643–661. <https://doi.org/10.1080/0907676X.2021.1953085>
- International Primary Care Respiratory Group. (2023). *How We Breathe* [Video]. YouTube. Retrieved from <https://www.youtube.com/watch?v=zrA3f5LzFeY>
- Jakobson, R. (1959). On linguistic aspects of translation. In R. A. Brower (Ed.), *On translation* (pp. 232–239). Harvard University Press. <https://doi.org/10.4159/harvard.9780674731615.c18>
- Jewitt, C., Bezemer, J., & O'Halloran, K. L. (2016). *Introducing multimodality*. Routledge. <https://doi.org/10.4324/9781315638027>
- Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice-Hall.
- Korning Zethsen, K. (2009). Intralingual translation: An attempt at description. *Meta*, 54(4), 795–812. <https://doi.org/10.7202/038904ar>
- Korning Zethsen, K. (2021). Intralingual translation. In Y. Gambier, & L. van Doorslaer (Eds.), *Handbook of translation studies* (Vol. 5, pp. 135–142). John Benjamins. <https://doi.org/10.1075/hts.5.int8>
- Korning Zethsen, K., & Hill-Madsen, A. (2016). Intralingual translation and its place within

translation studies: A theoretical discussion. *Meta*, 61(3), 692–708.
<https://doi.org/10.7202/1039225ar>

Kress, G., & van Leeuwen, T. (2020). *Reading images: The grammar of visual design* (3rd ed.). Routledge. <https://doi.org/10.4324/9781003099857>

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.

NHS. (2024). *Asthma* [Video]. Retrieved from <https://www.nhs.uk/conditions/asthma/>

Peruzzo, K., & Rocco, G. (2022). Exploring language simplification and intralingual translation: Insights, results and desiderata. *Rivista Internazionale di Tecnica della Traduzione*, 24, 7–21.

Rizzo, A. (2019). Museums as disseminators of niche knowledge: Universality in accessibility for all. *Journal of Audiovisual Translation*, 2(2), 92–136.
<https://doi.org/10.47476/jat.v2i2.93>

Note

Note 1. In this study, “intralingual subtitles” refers to the written reformulation or representation of spoken discourse within the same language, viewed from a translation-studies perspective. “Captions” is used primarily as an accessibility-related or platform-specific term and may also include speaker identification, sound effects, music, and other non-verbal auditory information. Although the two terms frequently overlap in current audiovisual practice, they are therefore not treated as entirely synonymous. The study retains “captions” when referring to the terminology adopted by the platforms themselves and uses “intralingual subtitling” for the broader mediating process under analysis.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>)