Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Whose Voice Counts? Synthetic-Voice Labour and Attributional Displacement on Douyin

Tianyi Gu

TianyiGu1992@163.com

Columbia University

Research Article

Keywords: AAC, attributional displacement, disability media, Douyin, platform labour, synthetic voice, recognised authorship

Posted Date: November 10th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10574479/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Whose Voice Counts? Synthetic-Voice Labour and Attributional Displacement on Douyin

Tianyi Gu

PhD Candidate

Columbia Journalism School, Columbia University

New York, NY, United States

Corresponding author: Tianyi Gu

Postal address: Columbia Journalism School, Columbia University, New York, NY, United States

Email: TianyiGu1992@163.com

ORCID: https://orcid.org/0009-0002-3271-8597

Running title: Synthetic-Voice Labour on Douyin

Article type: Original Research Article

Declarations

Funding: The author received no specific funding for this work.

Competing interests: The author declares no competing interests.

Author contribution: T.G. contributed to conceptualization, methodology, investigation, formal analysis, data curation, visualization, writing, review and editing, and project administration.

Ethics approval: Institutional ethics approval was obtained before fieldwork. The committee name and approval number must be inserted before submission.

Informed consent: Communication-accessible informed consent was obtained. Public availability was not treated as blanket consent.

Data availability: Identifying voice, video, account, and linkage data are restricted. Non-identifying instruments, coding rules, and audit materials are provided in the supplement.

Use of generative AI: Generative AI was used only for language editing and translation assistance; all empirical materials and interpretations were human-verified.

Whose Voice Counts? Synthetic-Voice Labour and Attributional Displacement on Douyin

Abstract

A smooth synthetic voice can make a long act of augmentative and alternative communication (AAC) sound immediate. This article examines how that audibility is produced and how audiences decide whether machine-mediated words, emotions and competence belong to the nonspeaking creator. Drawing on a six-month netnographic study conducted between September 2025 and February 2026, the analysis combines 14 adult creators, 487 public videos, 412 hours of platform observation, 423 purposively retained voice- and authorship-related comments/replies, and 21 communication-accessible interviews totalling 1,087 minutes. Rather than claiming a universal sequence, the findings identify a recursive configuration: creators compose through heterogeneous AAC and voice arrangements, format expression for platform genres, encounter audience attribution, and sometimes undertake verification labour. The article develops attributional displacement—the movement of presumed authorship, competence, or affect away from the creator toward a device, communication partner, AI system, or commercial intermediary when external involvement is interpreted as replacing rather than supporting the creator's role. This reframes platform accessibility as a problem of circulation, attribution, tempo and governance, not interface operation alone.

Keywords: *AAC; attributional displacement; disability media; Douyin; platform labour; synthetic voice; recognised authorship*

1. Introduction

The video appears to begin with a voice. A creator faces the camera while a synthetic narrator tells a story about work, intimacy, disability, humour or ordinary life. What audiences hear in thirty seconds may have taken hours to compose through eye-gaze typing, an AAC application, a phone keyboard, predictive text or repeated selections from a voice interface. The finished video begins after much of the labour has disappeared. Viewers hear fluency, while fieldnotes record waiting, selecting, correcting, rehearsing and deciding whether the composition process should be shown.

The disappearance of this labour reorganises recognition. Some viewers respond to the creator's argument or joke; others ask whether a parent wrote the script, whether the voice is 'really' theirs, or whether the person on screen understands what is being said. The central puzzle is therefore not simply whether technology enables speech. It is how technologically mediated expression is attributed when platform culture rewards speed, polish and affective immediacy. Audibility does not automatically produce recognised authorship.

This puzzle is consequential for nonspeaking and speech-disabled creators. The term nonspeaking is used here as a communication-centred rather than diagnosis-centred descriptor: participants could not rely on natural speech as their primary or consistently intelligible mode of communication and used AAC, text-to-speech, voice banking, eye-gaze input, delegated narration, captions or other non-oral modes in platform production. The article follows a communicative arrangement rather than a diagnostic category.

Douyin is a generative field because voice, body, captions, editing rhythm, recommendation systems, audience reaction and commerce converge in one environment. For nonspeaking creators,

these elements do not merely carry messages already formed elsewhere. They shape what can circulate: slow composition may be cut, a generic young voice may travel better, a brand may prefer a smoother tone, and an educational post may reveal eye-gaze typing to prove authorship.

Douyin matters here not as a generic short-video container but because the observed field combined compressed video pacing, integrated voice-and-caption production, creator-facing engagement metrics, MCN and brand coordination, and livestream commerce. Across fieldnotes and interviews, these elements made synthetic voice simultaneously a communicative interface, an editing object, a recognisable persona and a commercial asset. The platform-specific claim is therefore about how these elements were assembled in this field, not that all short-video platforms operate identically.

Existing scholarship explains parts of this scene but not the specific attribution problem. AAC research explains communicative competence, embodiment and voice personalisation; disability media scholarship shows that access is relational and shaped by ableist norms; platform research explains visibility, self-branding and precarity; and generative-AI research shows how models can accelerate or reshape expression. What remains under-theorised is the moment when a nonspeaking person becomes a public creator whose mediated utterance must circulate as content while remaining attributable to them as its author.

This article addresses that gap through a qualitative design linking what creators publish, how audiences respond and how creators explain production decisions. The evidence base is primary and study-generated: platform observations, interview transcripts, public video records, comment threads, fieldnotes, coding outputs and case matrices were produced and analysed by the research team. The analysis follows mediated utterances across composition, formatting, circulation, attributional judgement and creator response, while avoiding causal claims that the available platform traces cannot support.

The theoretical argument draws on Callon's concept of socio-technical agencements but shifts the analytic object from creator visibility to attributed authorship. A platform voice is assembled across the creator, AAC interface, predictive system, synthetic-voice library, communication partners, editing tools, audience expectations and commercial arrangements. The article makes four contributions. First, it distinguishes public audibility from recognised authorship. Second, it specifies synthetic-voice labour as the transformation of slow, multimodal composition into an apparently immediate vocal performance. Third, it develops attributional displacement to explain how words, feeling or competence can be reassigned from the creator to a device, partner, AI system or brand. Fourth, it theorises recognised authorship as a contingent achievement grounded in selection, veto, account control, disclosure and accountability.

2. Literature Review: From Communication Access to Synthetic-Voice Platform Labour

The review is organised as a longitudinal dialogue rather than a topical inventory. It traces a shift from AAC as 'communicative competence' (Light & McNaughton, 2014), to disability media as 'restricted access' (Ellcessor, 2016), to platform studies of the 'platformization of cultural production' (Nieborg & Poell, 2018), and finally to generative-AI debates about machine-mediated authorship. A condensed historical map is provided in Supplementary Table G6.

2.1 From functional access to communicative personhood

Modern AAC scholarship first established that people with complex communication needs can select symbols, construct messages and participate through modes other than natural speech. The Convention on the Rights of Persons with Disabilities codifies communication as including 'augmentative and alternative modes, means and formats of communication' (United Nations, 2006, Article 2). This rights-based premise remains crucial, but early functional accounts largely examined clinical, educational, service or family settings rather than public content, metrics, sponsorship or audience judgement.

Light and McNaughton's (2014) account of 'communicative competence' widened the analytic unit from device operation to linguistic, operational, social and strategic resources. The mobile era further normalised portable speech-generating technologies: McNaughton and Light (2013) described a 'mobile technology revolution', while Broomfield et al. (2024a, 2024b) foreground AAC users' experiences across changing lives. Voice synthesis research adds that intelligibility is also identity work. Hyppa-Martin et al. (2024) show how earlier speech-generating-device voices produced 'depersonalized SGD use'. Yet when a personalised or synthetic voice becomes a public creator persona, questions of ownership, editing authority and recognition are relocated to platform circulation.

2.2 From access to interdependence and disability media

Disability media and critical access studies challenge the idea that inclusion can be achieved by attaching a tool to an unchanged environment. Ellcessor's (2016) 'restricted access' and Alper's (2017) 'giving voice' shift attention from what a device permits to the conditions under which mediated utterance is heard as credible and consequential. Bennett et al. (2018) extend this move through 'interdependence as a frame', showing that disabled people often act through human and technological coordination. For nonspeaking creators, interdependence may enable filming, typing, translation or brand negotiation; it may also become editing, normalisation or gatekeeping.

This tradition also shows that authenticity is socially demanded rather than simply possessed. Bitman (2022) identifies 'constant able-bodied surveillance and disciplining', while Duffy and Hund (2019) theorise an 'authenticity bind' in social media self-branding. Synthetic voice intensifies this bind: polished speech can increase intelligibility while making authorship more suspect. The present article therefore extends access theory with a platform-labour vocabulary that explains why mediated expression becomes measurable, profitable and self-managed.

2.3 Platformisation, visibility and disabled creator labour

Platform studies recast cultural production as dependent on infrastructures of distribution, data extraction, monetisation and governance. Nieborg and Poell's (2018) 'platformization of cultural production' and Poell et al.'s (2021) account of platforms and cultural production show that cultural goods are continuously adjusted to platform markets. Creator-labour scholarship then identifies the insecurity behind entrepreneurial narratives: Duffy's (2017) 'aspirational work', Duffy's (2020) 'algorithmic precarity' and Duffy et al.'s (2021) 'nested precarities' explain how unstable recognition and income structure social media work. Abidin's (2016) 'visibility labour', Cotter's (2019) 'visibility game' and Bishop's (2019, 2020) work on algorithmic lore show how creators manufacture presence

under opaque conditions; Bucher (2018), Gillespie (2018) and Duffy and Meisner (2023) clarify that platform governance is neither neutral nor evenly experienced.

Disabled-creator research brings these problems closer to the present case. Choi et al. (2022) formulate an 'inclusive algorithmic experience'; Heung et al. (2024) show how disabled creators become vulnerable to ableist harassment; Lyu and Carroll (2024) identify BlindTokers' infrastructuring work; and Lyu et al. (2024) describe flourishing labour. Lu and Guo's (2026) Douyin study of Deaf creators is the closest precedent. It explains a visibility paradox in which exposure and livelihood opportunities coexist with hearing-centred commercial discipline, cultural commodification and constrained creative autonomy. The present study does not treat synthetic voice as merely another route to visibility; it asks how a mediated utterance is assigned to, or detached from, the person who composed it.

2.4 From visibility to attributed authorship: delimiting the contribution

The distinction is substantive rather than terminological. Lu and Guo (2026) centre culturally Deaf creators whose principal linguistic practice is sign language and examine how subtitles, AI dubbing, Signed Chinese and hearing 'coaches' shape visibility, cultural autonomy and commercial participation. This study centres nonspeaking and AAC-using creators for whom the contested object is the source of the utterance itself: whether the words were selected by the creator, generated by a system, written by a relative, scripted by a brand or understood by the person on screen. Its unit of analysis is therefore the voice arrangement and the attribution episode, not creator visibility alone.

This shift changes the theoretical question. Visibility research asks what creators must do to be seen and monetised; attributed-authorship research asks whose intention, competence and affect are recognised after technologically mediated speech circulates. The comparison is summarised in Supplementary Table G7. The contribution claimed here is accordingly narrower than a general platform paradox: it concerns the displacement and verification of authorship under conditions of communicative difference.

2.5 Generative AI, voice technologies and the unresolved authorship problem

Recent work integrates predictive and generative systems into AAC. Gaines and Vertanen (2025) adapt large language models for character-based AAC; Karmakar and Sinha (2024) automate text-to-AAC generation; Pal et al. (2024) connect personal narratives with conversational AI; and Benevento et al. (2025) and Frisch and Vertanen (2025) stress evaluation in real-world contexts of participation and agency. McNally et al. (2024) report disabled users describing ChatGPT as a communication assistant and conversational partner. These systems can support composition, but the analytical consequences differ across generic TTS, personalised voices, voice banking, predictive text, LLM assistance and delegated narration. They should not be collapsed into a single category of 'AI voice'.

AAC scholarship also cautions against separating devices from bodies and relationships. Benson-Goldberg et al. (2025) show that embodied resources remain necessary in speech-generating-device interactions, while Pradana (2025) describes AAC devices through experiences of embodiment and disembodiment. Voice banking research demonstrates that accent, age, naturalness and personalisation are identity questions rather than technical properties alone (Chen et al., 2023; Hyppa-Martin et al., 2024). Critical voice research further shows that sounding voice can be datafied and treated as evidence of identity by institutions and platforms (Leix Palumbo & Prey, 2024). These

insights make authorship attribution central: a mediated voice is simultaneously an interface output, bodily performance, identity claim and potentially valuable data object.

Communication research has long shown that audiences distinguish human, machine and institutional sources even when message content is held constant (Sundar & Nass, 2001). Human–machine communication and AI-mediated communication research further treats AI not simply as a channel but as a communicative actor or message-transforming intermediary (Guzman & Lewis, 2020; Hancock et al., 2020). Attributional displacement extends these traditions by examining a disability-specific epistemic asymmetry: technological or relational support is treated not merely as a source cue but as a reason to reassign the creator's intention, competence or feeling elsewhere.

Qualitative research with AAC users raises parallel epistemic questions. Keane and Kocsis (2025) diagnose missing methods and epistemic injustice, Noyes and Wilkinson (2024) centre adult AAC users' service experiences, and Temple and Young (2004) identify translation dilemmas. Agency cannot be studied if communication modes are regularised beyond recognition or if partner contributions are silently attributed to creators. The gap is therefore precise: existing traditions do not yet explain how a platformed synthetic voice becomes an access mechanism, authored performance, commercial resource and object of public source attribution at the same time.

3. Theoretical Framework

Callon's (2008) concept of socio-technical agencements provides a vocabulary for following mediated authorship without reducing it to either technology or individual choice. Agency is produced through arrangements of people, devices, competencies, rules, resources and calculative procedures. A synthetic-voice video is therefore not a transparent output of intention; it emerges through message selection, prediction, interface design, assistance, editing conventions, audience anticipation and platform metrics. The analytical unit is the arrangement that carries an utterance outward and the attributional judgement that follows.

This perspective avoids two errors. A technological-determinist account treats AAC or AI voice as empowerment or alienation in itself; an individualist account treats stylistic adaptation as unconstrained choice. Agencement analysis instead asks which capacities are enabled, who holds selection and veto rights, which intermediaries reshape expression, and how the resulting voice is publicly attributed. Predictive text may increase speed while privileging common phrasing; a partner may enable expression while becoming an aesthetic gatekeeper; and commercial metrics may provide income while narrowing communicative experimentation.

Callon's distinction between prosthetic and habilitated agency adds a normative dimension. In a prosthetic arrangement, a person is helped to operate within an environment whose standards remain intact. In a habilitative arrangement, the environment is reorganised to expand possible action. Applied to platform communication, a prosthetic orientation asks how nonspeaking creators can sound sufficiently rapid, fluent and ordinary for existing interfaces and audiences. A habilitative orientation asks how platforms might recognise plural tempos, visible composition, non-vocal modes, collaborative authorship and creator-controlled disclosure without treating these as deficiencies.

Synthetic-voice labour names the coordination and transformation work that converts communicative intention into a platform-circulating vocal performance. Its distinctiveness does not lie in claiming that composition, affective labour or relational labour are individually new. It lies in their convergence around a temporal and attributional transformation: slow, multimodal and often collaborative composition is compressed into speech that sounds immediate and singular.

Attributional displacement occurs when audiences move the presumed source of the utterance away from the creator and toward a device, communication partner, AI system, brand or other intermediary. The presence of technological or human mediation does not itself constitute attributional displacement; displacement occurs only when mediation is interpreted as replacing the creator as the presumed source of intention, competence or affect. Recognised authorship is the contingent social recognition that the creator selected, understood, authorised and remained accountable for the mediated expression. It does not imply that authorship can be objectively verified by audiences or that creators must publicly prove their competence. Here, recognised authorship describes how authorship becomes socially acknowledged through attribution processes; it is not a normative requirement that disabled creators disclose private production processes or publicly prove competence.

The concept is retained only where four conditions can be linked across more than one data type: mediated voice expands practical participation; the arrangement makes authorship or competence contestable; doubt is directed toward an external source; and the creator responds through explanation, process disclosure or decision-right claims. Counter-cases are necessary because accepted mediation, artistic convention and stable audience relations show that no technology produces the paradox by itself. Table 3 reports this theory-building test.

Figure 1 specifies this mechanism. It separates composition, voice arrangement, platform formatting and audience attribution; distinguishes recognised authorship from attributional displacement; and shows verification labour as a possible response rather than an inevitable final stage.

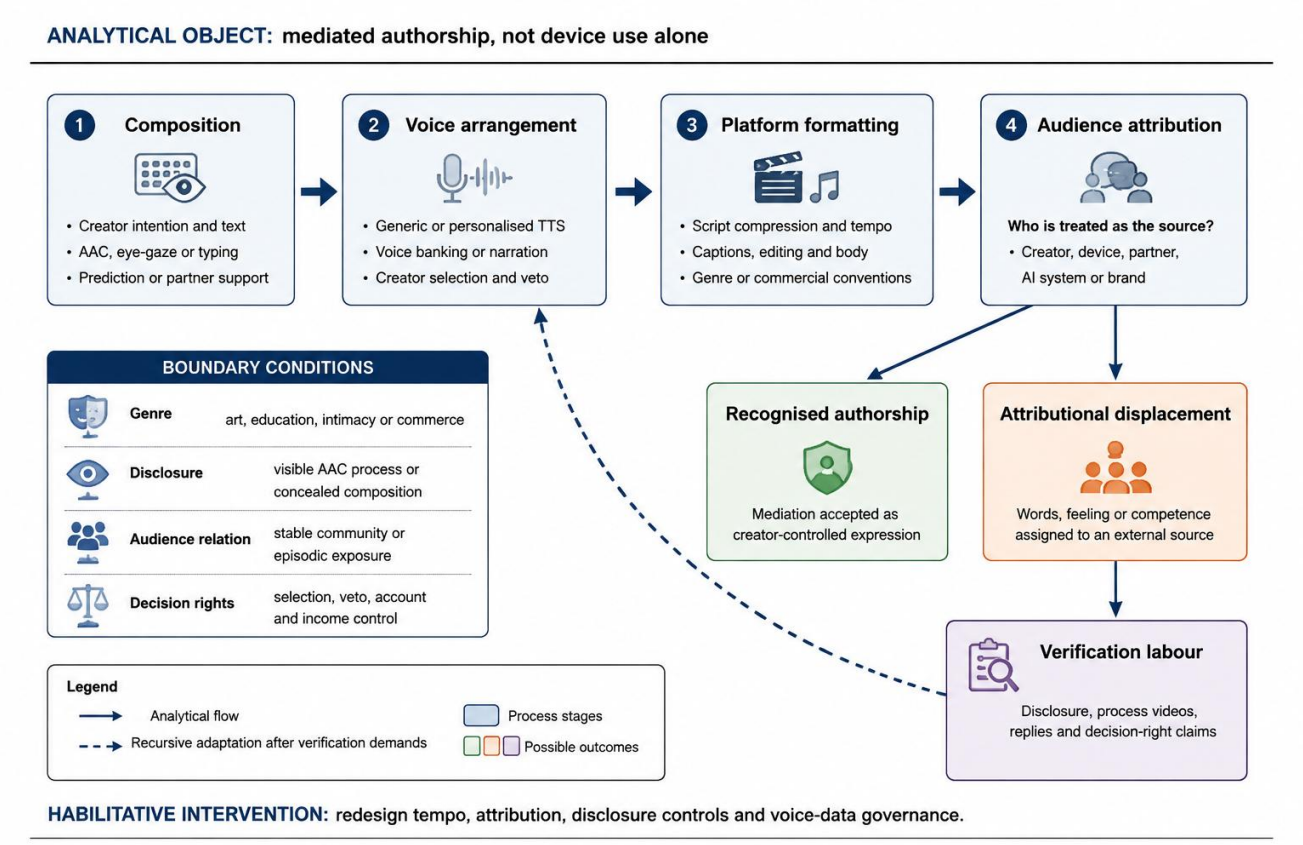


Figure 1. Analytical model of synthetic-voice labour, attributional displacement and recognised authorship.

Note. Solid arrows organise observed analytical relations; the dashed arrow represents recursive adaptation after verification demands. The model does not assert a universal causal sequence.

3.1 Research questions

RQ1. How do nonspeaking creators use different AAC and synthetic-voice arrangements to enter and reshape recognisable Douyin genres?

RQ2. How are selection, veto, editing, account control and commercial decisions distributed among creators, communication partners, technologies and intermediaries?

RQ3. Under what conditions is mediated speech attributed to the creator, and when are words, emotion or competence displaced toward a device, partner, AI system or brand?

RQ4. How do creators respond to attributional doubt, and which genre, disclosure and audience conditions reduce or intensify verification labour?

4. Methodology

This study adopts a qualitative netnographic design that connects public presentation, audience attribution and creator accounts without treating the materials as a causal panel. Videos show what circulated; fieldnotes situate platform events; comments and replies show how source and competence were publicly judged; and interviews explain production decisions, support relations and responses that finished posts often concealed. Triangulation is used to test and bound interpretations, not to transform qualitative materials into prevalence estimates. Table 1 summarises the empirical materials and their analytical functions.

Table 1. Empirical materials and analytical function

Data source	Completed scale	Analytical function
Netnographic observation	September 2025-February 2026; 412 hours	Contextualise platform norms, content changes, audience interaction and creator strategies
Video corpus	487 public videos from 14 selected creators	Map genres, voice arrangements, process visibility and public self-presentation
Accessible interviews	14 creators; 21 sessions; 1,087 minutes across text, captioned meetings, AAC output, asynchronous video/text and family-assisted communication with text confirmation	Explain motivations, production processes, decision rights, income and attribution experiences
Comments and creator replies	423 purposively retained items concerning voice, AAC, authorship, competence, commercial scripting, emotional sincerity or	Identify forms of attributional judgement and linked creator responses; not estimate comment

Data source	Completed scale	Analytical function
	creator response	prevalence
Engagement traces	Available likes, comments, shares, saves and follower ranges; newly posted videos captured at 14 days (+/- 2 days)	Provide contextual field traces only; not infer platform preference, causal effects or between-genre performance

Note. Counts locate the corpus and support qualitative mapping. Purposively retained comments do not provide a denominator for prevalence, and engagement traces are not used to infer platform reward or causal effects.

The study adapts netnography to a digitally mediated field (Kozinets, 2002; Pink et al., 2015; Postill & Pink, 2012). Douyin is treated not as a repository of videos but as a socio-technical field in which creators, followers, brands, interfaces and metrics interact. Fieldwork ran from September 2025 to February 2026. Two researchers observed selected accounts, new posts, profile changes, comments, creator replies, livestream announcements, sponsorship markers and relevant platform trends. Daily fieldnotes and weekly memos tracked changes in voice use, disclosure, audience judgement and creators' theories of visibility.

4.1 Sampling and corpus construction

Before formal sampling, the researchers conducted a feasibility scan using Chinese search terms related to speech disability and communication technologies. The scan identified 87 potentially relevant accounts; 31 met preliminary eligibility criteria; and 14 creators participated. Inclusion required adult status, self-identification as nonspeaking, speech-disabled or an AAC user, creator-authored or creator-selected synthetic or speech-generating communication in at least 30% of public videos during the sampling window, at least ten posts in the previous six months, personal participation in account or content decisions, and willingness to participate through a preferred communication mode. Purposive maximum-variation sampling sought diversity in follower range, gender, age, communication technology, content genre, location and commercialisation. Sample size was justified by information power (Malterud et al., 2016). The purpose of the sample was not statistical representation but mechanism-oriented theoretical development through comparison across communication arrangements, genres, and attribution contexts. The non-identifying participant profile is provided in Supplementary Table G1.

The final corpus comprised 487 public videos posted by the selected creators during fieldwork. The project did not use automated scraping. Metadata were recorded manually in a structured log, including posting date, topic, communication modality, available interaction counts, follower range and visible commercial disclosure. The 423 comments/replies analysed in depth were purposively retained after screening visible comment spaces for references to voice source, AAC process, authorship, competence, commercial scripting, emotional sincerity or creator response. Because the total visible-comment universe and platform ranking process were not consistently recoverable, these items are treated as an analytic corpus of attribution events, not a representative sample of audience opinion. No claim in the findings compares genre-specific prevalence of suspicion. Supplementary Table G2 reports corpus completeness and missingness.

4.2 Accessible interviewing and translation

Each creator was invited to one or two interviews, yielding 21 sessions and 1,087 minutes. The procedure was communication-accessible by design. Participants chose among WeChat text, Tencent Meeting with real-time captioning, AAC output, asynchronous video-plus-text or family-assisted oral communication with text confirmation. Questions were provided in advance, breaks were unrestricted, and response latency was not interpreted as uncertainty. Communication partners attended only at the participant's request, and their contributions were identified separately in transcripts. When partners assisted with communication access, they supported interactional logistics rather than substituting for participant responses. Partner explanations, interpretations, and independent statements were labelled separately and were not treated as participant-authored accounts. Interviews and posts were analysed first in Chinese or in the participant's original AAC text.

4.3 Analytical units, coding and evidentiary limits

The study used reflexive thematic analysis for interviews, fieldnotes, video narratives and creator self-explanations (Braun & Clarke, 2006, 2021). Analysis moved through familiarisation, open coding, memo writing, candidate-theme development, cross-case comparison, negative-case analysis and theoretical refinement. Structured video descriptors were kept separate from interpretive themes. Two researchers independently coded 120 videos (24.6% of the corpus), then refined the codebook. Agreement was substantial for primary narration modality (kappa = 0.81), AAC-process visibility (kappa = 0.78), perceived voice gender coding (kappa = 0.74) and commercial markers (kappa = 0.82). Interpretive themes were not subjected to mechanical agreement scoring.

Two linked analytical units were added during cross-case refinement. First, a voice-arrangement typology distinguished generic TTS, personalised synthetic voice, voice banking, AAC-integrated TTS and partner-supported or delegated narration according to text source, voice source, support role, creator selection and disclosure. These are configurations, not levels of technological sophistication. Second, an attribution episode linked a post, a comment or reply, a creator response and an interview or fieldnote only when the materials clearly referred to the same bounded event. The linked episodes support process interpretation but do not constitute a complete event census. Supplementary Tables G8-G10 document the typology, episode logic and denominator policy.

Engagement traces were retained to understand the field context and timing of creator decisions, but they were not standardised by follower count, paid promotion or recommendation exposure. Raw between-genre medians have therefore been removed from the main findings. The analysis does not claim that Douyin 'rewards' a voice type on the basis of these traces; commercial and visibility pressures are established through creator accounts, visible arrangements and bounded field episodes.

4.4 Ethics, identifiability and reflexivity

Institutional ethics approval was obtained before fieldwork; the approving body and reference number are withheld in the anonymised manuscript and should be supplied in the non-anonymised submission materials. Public availability was not treated as blanket consent. Account names, profile images, locations, searchable quotations and recognisable biographies were removed or paraphrased. Direct screenshots are not reproduced. Voice files, screen recordings and raw video are encrypted and

access-limited because they may be biometric or highly identifying. Participant profiles use broad ranges, and combinations of genre, age, technology and follower size were reviewed for reverse-identification risk. Researcher positionality was documented throughout, especially around fluency, speed, translation, family mediation and assumptions about what counts as independent authorship (Franzke et al., 2020).

In the anonymised manuscript, positionality is reported at the analytical rather than biographical level to avoid author identification. Reflexive attention focused on how assumptions about fluency, speed, translation, family mediation and solitary production could shape interpretation. The analysis therefore treated these standards as culturally and institutionally situated rather than neutral indicators of competence or authorship. Supplementary Table H1 summarises the qualitative accountability materials and audit trail.

5. Findings and Analysis

The findings are organised around a recursive configuration rather than a linear stage model. Each section follows the same evidentiary rhythm: thematic claim, field context, participant quotation or linked video/comment evidence, interpretation and boundary condition. Figure 2 shows how communicative access, platform legibility, attributional judgement and creator response can feed back into one another. The arrows organise recurring relations across cases; they do not claim that every creator follows the same chronology.

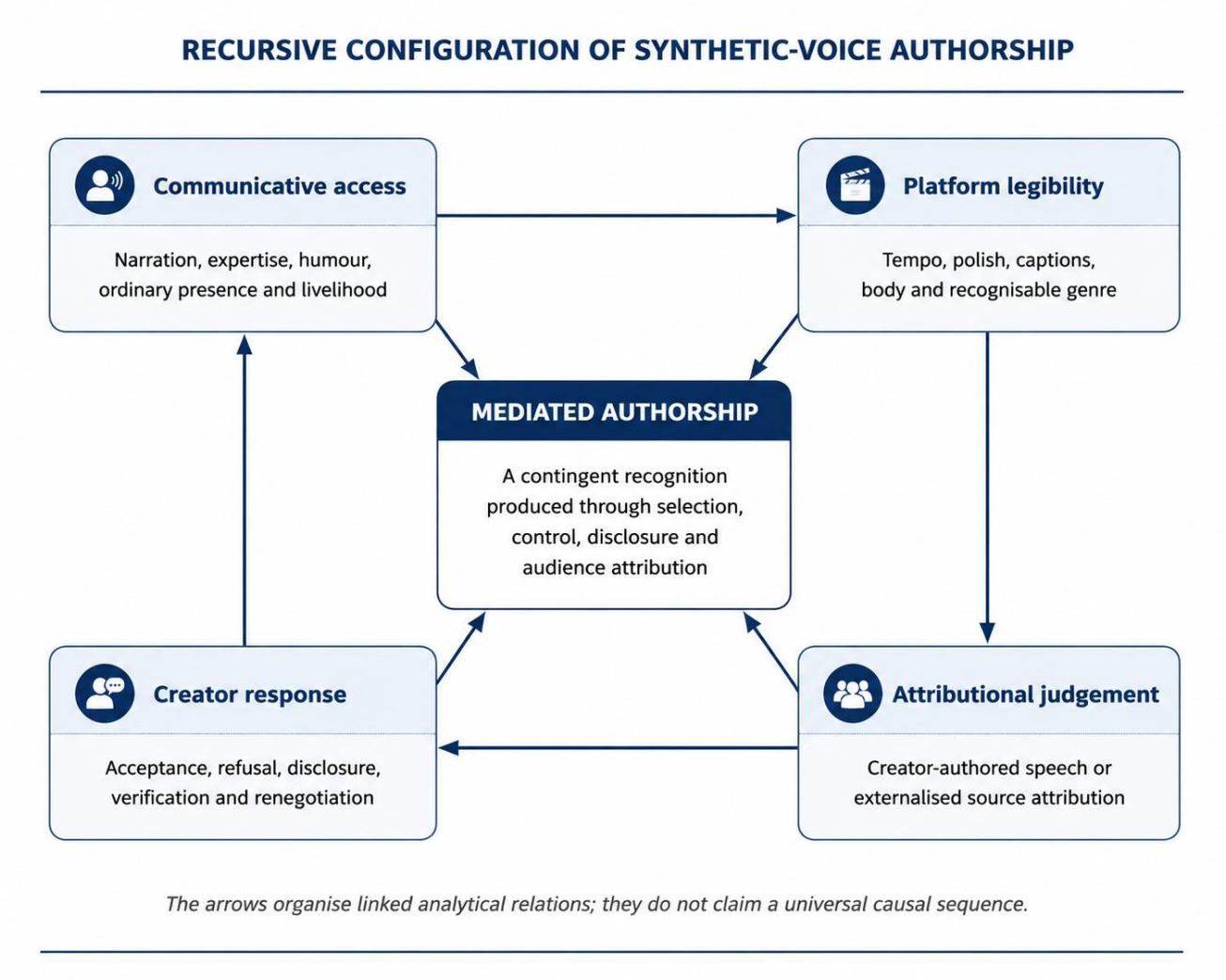


Figure 2. Recursive configuration of synthetic-voice authorship on Douyin.

Note. The configuration distinguishes linked analytical relations from causal or universally ordered stages. Accepted authorship and attributional displacement can coexist across genres and within the same account.

5.1 What nonspeaking creators produce on Douyin

The corpus shows that nonspeaking creators use synthetic voice not only to compensate for absent or unreliable speech but to enter recognisable platform genres: daily-life narration, disability explanation, humour, assistive-technology education, commercial promotion and intimate content. Synthetic voice opened spaces for ordinary presence, expertise, humour, income and relational expression, yet each genre posed different demands concerning tempo, disclosure, affect and authorship. Table 2 maps corpus composition and interpretive stakes; it does not compare engagement performance.

Table 2. Corpus composition, voice arrangements and authorship stakes

Primary theme	Common voice arrangement	Primary communicative and authorship stake
Daily-life narration	Generic TTS + captions	Ordinary presence; making everyday experience narratable without requiring disclosure in every post
Disability explanation and advocacy	Personalised or voice-banked speech + visible AAC	Explaining mediation and asserting authorship, competence and communication rights
Humour/comedy	Generic TTS + fast tempo	Using synthetic rhythm as comic timing and a repeatable platform persona
Assistive-technology education	AAC/TTS demonstration	Making mediation visible and teachable; shifting attention from deficit to practice
Commercial promotion	Polished TTS + commercial aesthetics	Negotiating brand scripts, pace, voice style and proof of creator control
Intimate/emotional content	Personalised TTS + affective tone	Making synthetic prosody carry vulnerability while retaining emotional authorship
Other/mixed	Mixed arrangements	Preserving cases that cross genres or resist stable classification

Note. Counts describe corpus composition and primary thematic classification. Categories are used for qualitative mapping and do not estimate audience prevalence, engagement performance or causal effects.

Daily-life narration was the largest category, followed by disability explanation and advocacy. The distribution indicates that synthetic voice supported both ordinary presence and explicitly disability-related communication. Genre also organised what was made visible. Advocacy and educational posts more often displayed AAC devices, typing or eye-gaze selection, while many humorous and commercial posts foregrounded finished fluency. Because exposure and follower scale were not standardised, this contrast is interpreted as a difference in presentation strategy, not evidence that the platform rewarded one arrangement over another.

P03 described this genre-specific calibration as a conflict between recognition and proof: 'For advocacy videos, I show my eye-gaze board because I want people to know this is really me. But for brand collaborations, the brand asked me to use the smooth voice and cut out the typing part; they said it was more professional.' A fieldnote from a P06 video shows the alternative strategy: a split frame displayed AAC text being composed word by word while P06's face reacted as each word appeared. The caption read, 'This is how I speak. It takes time. Thank you for waiting.' These examples show both adaptation and resistance: creators learned platform conventions, but they also used disclosure, slowness and embodied presence to contest the assumption that fluent voice is the only credible voice.

5.2 Voice arrangements and the making of platform legibility

Platform legibility was not produced by a device alone. Machine-mediated speech had to become intelligible, emotionally familiar and temporally compatible with short-video circulation. This involved negotiating accessibility, creative style and perceived platform norms around speed, retention, affect and repeatable persona. The same creator could use different arrangements across advocacy, commerce, intimacy or art.

Synthetic-voice labour included selecting messages, accepting or rejecting predictions, compressing long ideas, choosing voice styles, correcting pronunciation, coordinating captions, synchronising gesture and deciding whether to reveal composition. All 14 participants discussed shortening or simplifying expression for video pacing, and 12 discussed selecting voices by age, gender, accent, speed or emotional tone. In the corpus, 198 videos (40.7%) visibly displayed the communication process, while 289 (59.3%) did not. These figures show that much communicative work was absent from the final frame; they do not by themselves establish why a post was recommended or how audiences interpreted it. Supplementary Table G3 summarises the dimensions of synthetic-voice labour and their evidence base.

The technology typology clarifies what 'synthetic voice' covers. Generic TTS separated the creator's text from a shared platform voice; personalised synthesis and voice banking connected output more directly to a chosen or historically embodied vocal identity; AAC-integrated TTS linked alternative input to vocal output; and partner-supported narration added an explicit human intermediary. Authorship questions therefore differed by arrangement. The relevant issue was not whether AI was present in the abstract, but who selected the text and voice, who could reject changes, how support was disclosed, and which source audiences presumed. Supplementary Table G8 reports the typology.

This time compression was not a researcher inference. P01 explained, 'For a 30-second video, I sometimes write for two hours. I have to compress what I want to say into a few sentences, otherwise viewers swipe away.' P02 described voice selection in similarly practical terms: 'I tried seven different voices. At first I used a mechanical one. Later I found a younger and faster voice, and followers said, this sounds like you. So I kept using it.' The first quotation shows constraint; the second shows agency. Together, they show why synthetic-voice labour is strategic adaptation rather than passive technological use.

The embodied dimension was also central. P14 said, 'When I record a video, I have to listen to the rhythm of the synthetic voice, and my face has to move with it. If the voice laughs, I have to smile; if the voice sighs, I have to frown. It takes practice.' Synthetic voice does not replace the body; the body is recruited to authenticate the voice, supply affect that the system cannot deliver and make a mediated utterance appear socially present.

5.3 Attributional displacement, authorship scrutiny and accepted mediation

Within the retained attribution episodes, audience judgement sometimes displaced authorship away from the creator. Comments questioned whether the creator wrote the words, understood the message, controlled the account or genuinely felt the emotion carried by a synthetic voice. Attributional displacement names this change in presumed source: the creator moves from recognised speaker to the person for whom a device, parent, partner, AI system or commercial script is assumed to speak.

Five forms of scrutiny recurred in the retained attribution corpus: technological suspicion, family-authorship suspicion, competence doubt, commercial suspicion and emotional-authenticity doubt. These categories describe the forms encountered; they do not estimate their prevalence among all comments. Counter-cases of acceptance, creativity and stable community recognition were analysed alongside scrutiny. Linked episodes suggest that disclosure, technical demonstration, artistic framing and established audience relations could make mediation more intelligible, while commercial polish or visible partner involvement could provide occasions for external attribution. Supplementary Table G4 provides the classification.

Technological suspicion appeared repeatedly in the retained advocacy-related attribution records. P03 reflected: 'They do not really want to know the technical details. They want to confirm whether I am really disabled, as if I have to perform it for them to count.' Family-authorship suspicion appeared when a partner was visible. P06 explained, 'They see someone next to me and think it is not me. But no one asks a person typing on a keyboard, Did you actually type that?' Competence doubt linked speech difference to cognitive authority; P07 said, 'What hurts most is not that they say I cannot speak, but that they say I do not understand. They confuse my body with my mind.'

Commercial and emotional episodes made the attribution problem especially explicit in the linked records. P08 said of a brand collaboration, 'The thirty seconds viewers saw took me a whole morning. I just wanted to make good content; I was not deceiving anyone.' P10 located emotional authorship in the words she selected and the embodied work surrounding them, rather than in synthetic prosody alone. A counter-case qualifies the pattern: P13's art and music videos attracted comments about creativity rather than authorship. Synthetic voice alone therefore does not produce doubt; genre expectations, disclosure and ableist assumptions shape whether mediation is accepted or externalised.

5.4 Decision rights and recognised authorship

Family members, partners, therapists, editors, MCNs and brand contacts helped some synthetic voices become publishable. Their involvement does not automatically reduce agency. Recognised authorship depends on how decision rights are distributed: who proposes, selects, rejects, controls account access, communicates with brands, receives income and answers for the final post. This reframes independence as meaningful control rather than the absence of support. The distribution of decision rights across creators, partners, technologies and intermediaries is summarised in Supplementary Table G5.

P06's mother assists with selecting topics, filming and editing. P06 explained: 'My mom helps me press the keys because my hands are hard to control. But every word is something I have thought through. She does not choose words for me; she is just my hands.' By contrast, P08 described MCN involvement: 'They tell me this sentence is too long and Douyin audiences do not have patience. Sometimes they suggest a more explosive opening.' P06's 'just my hands' describes assistance that carries intention outward; P08's 'explosive opening' describes platform-oriented shaping. The difference lies not in whether another actor is present but in selection, veto and accountability. Supplementary Table G5 reports the decision-rights matrix.

All 14 participants reported retaining password access and formal final-posting authority. Formal account control nevertheless did not eliminate economic, relational or stylistic constraints: a creator could retain the final click while still facing meaningful costs for refusing a brand, MCN or family-preferred form.

5.5 Livelihood, commercial adaptation and vocal norms

For creators excluded from conventional employment or needing flexible work, platform production could provide income, recognition, peer connection and a public role. Commercial participation also introduced requests for faster, smoother or more conventional vocal styles. The evidence for this claim comes from creator accounts, visible brand arrangements and decision-right records—not from raw engagement differences between content genres.

Commercialisation was widespread but uneven. Eleven creators had brand partnerships, six displayed product links, three had livestream-commerce experience and eight received tipping or donation income. P02 linked income to speed: 'After my followers passed 100,000, brands started contacting me... In livestream commerce I use the fastest TTS voice because I can introduce more products.' P08 described a similar shift: 'I realised I had not used my slower voice for a long time—not because I could not, but because slow is not valuable on Douyin.' P11 provides a counter-case: 'If it is slow, it is slow. People have become used to it.' Commercial norms were influential but not total; autonomy depended on whether creators could refuse profitable voice expectations without losing meaningful participation.

5.6 A bounded configuration: the voice-authenticity paradox

The findings support a bounded voice-authenticity paradox rather than a universal sequence. Synthetic voice can expand narration, expertise, humour, advocacy and livelihood. Under some arrangements, however, the compression and polish that make expression circulate also increase ambiguity about who selected, understood or authorised it. Attributional displacement converts that ambiguity into an external source claim; verification labour is one possible creator response.

The voice-authenticity paradox is retained because three evidentiary conditions are present: synthetic voice increased platform legibility and practical participation; this legibility was linked to reception, livelihood or community recognition; and polished arrangements generated authorship scrutiny, identity constraint or verification labour. Table 3 provides a compressed theory-building check. It includes counter-cases so that the concept is bounded rather than universalised.

Table 3. Bounded evidence test for the voice-authenticity configuration

Analytical condition	Supporting evidence	Boundary or negative evidence	Interpretation
Mediated voice expands practical participation	All 14 creators described expanded expression; observed posts included narration, humour, education, commerce and intimacy	Some creators used slower or non-commercial arrangements without seeking maximum circulation	Access is established at the level of communicative possibility, not engagement causality
Authorship becomes contestable	Comments/replies and interviews questioned text source, competence, family authorship, commercial scripting or emotional sincerity	P13 art content and P11 stable non-commercial audiences showed limited scrutiny	Contestability is conditional on genre, audience expectation and ableist assumptions
Attribution shifts to an external source	Linked episodes assigned words or feeling to AI/TTS, a family member, partner, MCN or brand	Technical disclosure sometimes made mediation ordinary rather than suspicious	Attributional displacement, not mediation alone, is the mechanism
Creator undertakes verification labour	P03/P05/P06/P08 used replies, process videos, behind-the-scenes evidence or decision-right explanations	Some creators refused repeated proof or relied on established community recognition	Verification is a contingent response, not an inevitable final stage
Overall judgement	Evidence converges across interviews, videos, comments/replies and fieldnotes	No representative comment denominator; no causal platform-exposure measure	Retain as a bounded configuration and report evidentiary limits

Note. The test evaluates qualitative convergence and boundary conditions. It does not estimate audience prevalence, transition probabilities or causal effects; engagement traces are excluded because follower scale, exposure and promotion were not standardised.

P05 condensed the double bind: 'Every time someone says, you did not write this, right, I have to reply that I wrote it and that I use AAC. Sometimes I have to make another video just to show the typing process. It is tiring, but if I do not do it, the comments keep asking.' The public burden is therefore not simply to speak through technology, but to make the selection and production of that speech visible when authorship is displaced.

5.7 Linked attribution episodes: evidence for recursion, not a universal sequence

To examine process without converting the corpus into a causal panel, the analysis linked materials only when a post, comment or reply, creator response and interview or fieldnote clearly referred to the same bounded episode. P03 responded to AI-voice suspicion with a process video showing AAC selection. P06 answered family-authorship suspicion by distinguishing her mother's

physical assistance from her own word selection. P08 answered commercial-scripting suspicion with behind-the-scenes evidence of extended AAC writing. P05 repeatedly replied to authorship questions and sometimes produced typing-process videos. These records support a recurring movement from external attribution to verification, but they do not establish how frequently that movement occurred across all posts.

The counter-case is equally important. P13's artistic voice was treated as a creative form, so mediation did not require repeated proof. P11's audience had become accustomed to a slower voice. Together, the linked episodes show that creator response feeds back into later disclosure and presentation decisions, while genre and audience familiarity can interrupt the proof demand altogether. Supplementary Table G9 presents the bounded episode records and their analytic limits.

6. Discussion

6.1 From visibility to attribution

The empirical story changes the object of accessibility. Conventional questions ask whether a person can enter, perceive, operate or communicate through a system. Creator platforms add an attribution problem: communication must circulate, be interpreted and be recognised as the creator's authored expression. Platform accessibility therefore includes communicative tempo, audience intelligibility, source attribution and the right to be treated as a competent author of technologically mediated speech.

6.2 Attributional displacement and recognised authorship

This contribution is distinct from the closest Douyin precedent. Lu and Guo (2026) show how Deaf creators' visibility and creative autonomy are shaped by hearing-centred platform and commercial norms. The present study addresses a different communicative and epistemic problem: nonspeaking creators may become audible while the source of their words, emotions or competence is reassigned elsewhere. Visibility concerns whether creators enter circulation; attribution concerns whether circulation still counts as their authored speech. The two processes can interact, but they are not interchangeable.

6.3 Synthetic-voice labour as temporal and evidentiary transformation

Attributional displacement specifies the mechanism. A message may be selected by a creator, predicted by a system, corrected by a partner, vocalised by a commercial service and promoted by an algorithm. Distribution does not dissolve authorship, but it makes source attribution contestable. Audiences can treat technological or relational support as evidence that the creator neither composed nor understood the utterance. Recognised authorship answers this problem through meaningful control: selection, veto, account authority, consent to support, responsibility for publication and the capacity to explain or refuse changes. This extends distributed agency into a more precise account of distributed yet accountable authorship. Used analytically, recognised authorship describes how recognition is organised; it does not prescribe disclosure or public proof as conditions of legitimate authorship.

Synthetic-voice labour also refines platform-labour scholarship. Its analytical value is not a catalogue of already named labour forms. It identifies a specific transformation in which hours of composition, device selection, voice adjustment, caption coordination and bodily synchronisation are compressed into an apparently effortless vocal object. When that compression hides process, creators may later be required to reverse it by displaying typing, AAC selection or behind-the-scenes work. Labour therefore becomes both production work and evidence of authorship.

6.4 Design and governance implications

The findings support four practice principles. Platforms should recognise plural communicative tempos through multimodal replies, creator-controlled delays and AAC-aware disclosure options that do not automatically reduce distribution. Commercial intermediaries should use accessible agreements that specify topic, wording, voice style, posting timing, veto rights and income distribution. AAC and voice-synthesis designers should support author-controlled provenance and disclosure while protecting voice data, rather than defining success only through naturalness and speed. Finally, platforms should not make disabled creators repeatedly prove competence through public disclosure; governance should place responsibility on systems and intermediaries to make support relations legible without converting disability into a verification performance.

7. Scope and Limitations

A Douyin-centred study cannot represent all Chinese nonspeaking people, platform users or AAC practices. Public creators are likely to have more resources, support, confidence or technical access than people excluded from content production. Recruitment through public profiles may underrepresent private or small-community communication networks. The sample is heterogeneous in diagnosis and communication history, which strengthens communication-centred comparison but limits disorder-specific conclusions.

Engagement metrics are unstable platform traces and were not standardised by follower count, paid promotion or recommendation exposure. They are therefore excluded from inferential comparisons. The 423 retained comments/replies were purposively selected for relevance to voice and authorship; without a complete and stable comment denominator, the study cannot estimate how common suspicion was. Linked attribution episodes support bounded process interpretation but not a universal chronology or causal pathway. Interviews may also be shaped by fatigue, partner presence and the researchers' interpretation of nonstandard written forms. Accessible procedures, participant review and reflexive documentation reduce but do not eliminate these asymmetries.

The concepts also require calibration. Synthetic-voice labour should not imply that all machine-mediated speech is burdensome, and attributional displacement should not romanticise unmediated natural speech as the only authentic form. Many participants valued synthetic voice as an authentic extension of self, and some audiences accepted it as an ordinary creative persona. The contribution lies in locating constraint in specific arrangements of attribution, disclosure and decision rights rather than in technology or disability itself.

8. Conclusion

This article began with a nonspeaking creator who appears on Douyin through a fluent synthetic voice and an audience that decides where that voice comes from. Following mediated utterances across composition, formatting, circulation and response shows that synthetic speech is more than a technical solution. It is a channel, aesthetic, labour process, identity claim, commercial resource, data object and site of source attribution.

Synthetic-voice labour names the work through which slow, multimodal composition becomes immediate narration. Attributional displacement explains why the resulting words, feeling or competence may be reassigned to a device, partner, AI system or commercial intermediary. Recognised authorship, in turn, depends on meaningful selection, veto, control and accountability rather than unsupported ideals of solitary production. The analysis therefore reframes authenticity as accountable selection and embodied authorship rather than vocal naturalness. The task for platforms is not simply to make synthetic voices sound more natural, but to make mediated authorship socially, commercially and institutionally count.

References

Abidin, C. (2016). Visibility labour: Engaging with influencers' fashion brands and #OOTD advertorial campaigns on Instagram. Media International Australia, 161(1), 86–100. https://doi.org/10.1177/1329878X16665177

Alper, M. (2017). Giving voice: Mobile communication, disability, and inequality. MIT Press.

Benevento, A. D., Ciulla, G., & Merlo, G. (2025). Future technologies in alternative and augmented communication: A scoping review of innovations. Frontiers in Communication, 10, Article 1607531. https://doi.org/10.3389/fcomm.2025.1607531

Bennett, C. L., Brady, E., & Branham, S. M. (2018). Interdependence as a frame for assistive technology research and design. In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility (pp. 161–173). ACM. https://doi.org/10.1145/3234695.3236348

Benson-Goldberg, S., Gullion, L., & Erickson, K. (2025). The role of the body in SGD-mediated interactions. Augmentative and Alternative Communication, 41(1), 1–16. https://doi.org/10.1080/07434618.2024.2398436

Bishop, S. (2019). Managing visibility on YouTube through algorithmic gossip. New Media & Society, 21(11–12), 2589–2606. https://doi.org/10.1177/1461444819854731

Bishop, S. (2020). Algorithmic experts: Selling algorithmic lore on YouTube. Social Media + Society, 6(1), 1–11. https://doi.org/10.1177/2056305119897323

Bitman, N. (2022). 'Authentic' digital inclusion? Dis/ability performances on social media by users with concealable communicative disabilities. New Media & Society, 24(2), 401–419. https://doi.org/10.1177/14614448211063183

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qualitative Research in Psychology, 18(3), 328–352. https://doi.org/10.1080/14780887.2020.1769238

Broomfield, K., Harrop, D., Jones, G. L., Sage, K., & Judge, S. (2024a). A qualitative evidence synthesis of the experiences and perspectives of communicating using augmentative and alternative communication (AAC). Disability and Rehabilitation: Assistive Technology, 19(5), 1802–1816. https://doi.org/10.1080/17483107.2022.2105961

Broomfield, K., Judge, S., Sage, K., Jones, G. L., & James, D. (2024b). Using longitudinal qualitative research to explore the experience of receiving and using augmentative and alternative communication. International Journal of Language & Communication Disorders, 59(3), 1043–1065. https://doi.org/10.1111/1460-6984.12981

Bucher, T. (2018). If...then: Algorithmic power and politics. Oxford University Press.

Callon, M. (2008). Economic markets and the rise of interactive agencements: From prosthetic agencies to habilitated agencies. In T. Pinch & R. Swedberg (Eds.), Living in a material world: Economic sociology meets science and technology studies (pp. 29–56). MIT Press.

Chen, M., Hyppa-Martin, J., Bunnell, H. T., Lilley, J., Foo, C., Tan, H. W., & Lim, W. S. (2023). Voice banking to support individuals who use speech-generating devices: Development and evaluation of Singaporean-accented English synthetic voices and a Singapore Colloquial English recording inventory. Augmentative and Alternative Communication, 39(4), 208–218. https://doi.org/10.1080/07434618.2023.2181213

Choi, D., Lee, U., & Hong, H. (2022). 'It's not wrong, but I'm quite disappointed': Toward an inclusive algorithmic experience for content creators with disabilities. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (Article 593, pp. 1–19). ACM. https://doi.org/10.1145/3491102.3517574

Cotter, K. (2019). Playing the visibility game: How digital influencers and algorithms negotiate influence on Instagram. New Media & Society, 21(4), 895–913. https://doi.org/10.1177/1461444818815684

Duffy, B. E. (2017). (Not) getting paid to do what you love: Gender, social media, and aspirational work. Yale University Press.

Duffy, B. E. (2020). Algorithmic precarity in cultural work. Communication and the Public, 5(3–4), 103–107. https://doi.org/10.1177/2057047320959855

Duffy, B. E., & Hund, E. (2019). Gendered visibility on social media: Navigating Instagram's authenticity bind. International Journal of Communication, 13, 4983–5002.

Duffy, B. E., & Meisner, C. (2023). Platform governance at the margins: Social media creators' experiences with algorithmic (in)visibility. Media, Culture & Society, 45(2), 285–304. https://doi.org/10.1177/01634437221111923

Duffy, B. E., Pinch, A., Sannon, S., & Sawey, M. (2021). The nested precarities of creative labor on social media. Social Media + Society, 7(2), 1–12. https://doi.org/10.1177/20563051211021368

Ellcessor, E. (2016). Restricted access: Media, disability, and the politics of participation. New York University Press.

Franzke, A. S., Bechmann, A., Zimmer, M., & Ess, C. (2020). Internet research: Ethical guidelines 3.0. Association of Internet Researchers. https://aoir.org/reports/ethics3.pdf

Frisch, B., & Vertanen, K. (2025). Designing AAC for use in social and community contexts: A scoping review. Augmentative and Alternative Communication. Advance online publication. https://doi.org/10.1080/07434618.2025.2558851

Gaines, D., & Vertanen, K. (2025). Adapting large language models for character-based augmentative and alternative communication. In Findings of the Association for Computational Linguistics: EMNLP 2025 (pp. 15273–15291). Association for Computational Linguistics. https://doi.org/10.18653/v1/2025.findings-emnlp.826

Gillespie, T. (2018). Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media. Yale University Press.

Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A Human–Machine Communication research agenda. New Media & Society, 22(1), 70–86. https://doi.org/10.1177/1461444819858691

Hancock, J. T., Naaman, M., & Levy, K. (2020). AI-mediated communication: Definition, research agenda, and ethical considerations. Journal of Computer-Mediated Communication, 25(1), 89–100. https://doi.org/10.1093/jcmc/zmz022

Heung, S., Jiang, L., Azenkot, S., & Vashistha, A. (2024). 'Vulnerable, victimized, and objectified': Understanding ableist hate and harassment experienced by disabled content creators on social media. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (Article 744, pp. 1–19). ACM. https://doi.org/10.1145/3613904.3641949

Hyppa-Martin, J. K., Lilley, J. S., Chen, M., Friese, J., Schmidt, C., & Bunnell, H. T. (2024). A large-scale comparison of two voice synthesis techniques on intelligibility, naturalness, preferences, and attitudes toward voices banked by individuals with amyotrophic lateral sclerosis. Augmentative and Alternative Communication, 40(1), 31–45. https://doi.org/10.1080/07434618.2023.2262032

Karmakar, P., & Sinha, M. (2024). Aiding non-verbal communication: A bidirectional language agnostic framework for automating text to AAC generation. In Proceedings of the 21st International Conference on Natural Language Processing (ICON) (pp. 324–331). NLP Association of India. https://aclanthology.org/2024.icon-1.37/

Keane, A., & Kocsis, J. (2025). Missing methods and epistemic injustice: A scoping review of qualitative research with AAC users. Disability & Society, 40(9), 2564–2587. https://doi.org/10.1080/09687599.2025.2458005

Kozinets, R. V. (2002). The field behind the screen: Using netnography for marketing research in online communities. Journal of Marketing Research, 39(1), 61–72. https://doi.org/10.1509/jmkr.39.1.61.18935

Leix Palumbo, D., & Prey, R. (2024). Sounding out voice biometrics: Comparing and contrasting how the state and the private sector determine identity through voice. Big Data & Society, 11(4), 1–15. https://doi.org/10.1177/20539517241297889

Light, J., & McNaughton, D. (2014). Communicative competence for individuals who require augmentative and alternative communication: A new definition for a new era of

communication? Augmentative and Alternative Communication, 30(1), 1–18. https://doi.org/10.3109/07434618.2014.885080

Lu, P., & Guo, Z. (2026). The sound of silence: Chinese Deaf creators' self-presentation, labour practices and the visibility paradox on Douyin. New Media & Society, 28(6), 2652–2672. https://doi.org/10.1177/14614448251338497

Lyu, Y., & Carroll, J. M. (2024). 'Because some sighted people, they don't know what the heck you're talking about': A study of blind TikTokers' infrastructuring work to build independence. Proceedings of the ACM on Human-Computer Interaction, 8(CSCW1), Article 20. https://doi.org/10.1145/3637297

Lyu, Y., Cai, J., Dosono, B., Yadav, D., & Carroll, J. M. (2024). 'I upload ... all types of different things to say the world of blindness is more than what they think it is': A study of blind TikTokers' identity work from a flourishing perspective. Proceedings of the ACM on Human-Computer Interaction, 8(CSCW2), Article 474. https://doi.org/10.1145/3687013

Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample size in qualitative interview studies: Guided by information power. Qualitative Health Research, 26(13), 1753–1760. https://doi.org/10.1177/1049732315617444

McNally, K., Wright, K., Goldkind, L., Kattari, S. K., & Victor, B. G. (2024). Disability expertise and large language models: A qualitative study of autistic TikTok creators' use of ChatGPT. Social Media + Society, 10(3), 1–10. https://doi.org/10.1177/20563051241279549

McNaughton, D., & Light, J. (2013). The iPad and mobile technology revolution: Benefits and challenges for individuals who require augmentative and alternative communication. Augmentative and Alternative Communication, 29(2), 107–116. https://doi.org/10.3109/07434618.2013.784930

Nieborg, D. B., & Poell, T. (2018). The platformization of cultural production: Theorizing the contingent cultural commodity. New Media & Society, 20(11), 4275–4292. https://doi.org/10.1177/1461444818769694

Noyes, A. M., & Wilkinson, K. M. (2024). A qualitative pilot study of adult AAC users' experiences related to accessing and receiving mental health services. Augmentative and Alternative Communication, 40(3), 238–252. https://doi.org/10.1080/07434618.2023.2294734

Pal, S., Das, S., Srihari, R., Higginborham, J., & Bizovi, J. (2024). Empowering AAC users: A systematic integration of personal narratives with conversational AI. In Proceedings of the 1st Workshop on Customizable NLP: Progress and Challenges in Customizing NLP for a Domain, Application, Group, or Individual (CustomNLP4U) (pp. 12–25). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.customnlp4u-1.2

Pink, S., Horst, H., Postill, J., Hjorth, L., Lewis, T., & Tacchi, J. (2015). Digital ethnography: Principles and practice. SAGE.

Poell, T., Nieborg, D., & Duffy, B. E. (2021). Platforms and cultural production. Polity.

Postill, J., & Pink, S. (2012). Social media ethnography: The digital researcher in a messy web. Media International Australia, 145(1), 123–134. https://doi.org/10.1177/1329878X1214500114

Pradana, W. A. (2025). In dialogue with the body: A phenomenological exploration of the interrelationship between people who use AAC and their AAC devices. Augmentative and Alternative Communication, 41(2), 99–113. https://doi.org/10.1080/07434618.2024.2407792

Sundar, S. S., & Nass, C. (2001). Conceptualizing sources in online news. Journal of Communication, 51(1), 52–72. https://doi.org/10.1111/j.1460-2466.2001.tb02872.x

Temple, B., & Young, A. (2004). Qualitative research and translation dilemmas. Qualitative Research, 4(2), 161–178. https://doi.org/10.1177/1468794104044430

United Nations. (2006). Convention on the rights of persons with disabilities. United Nations. https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-rights-persons-disabilities

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- WhoseVoiceCountsSupplementaryAppendices.docx