



## PODCASTS AND VLOGS AS A CREATIVE ENVIRONMENT FOR DEVELOPING SPEECH SKILLS

<https://doi.org/10.5281/zenodo.20698977>

**Kamolova Umida Bektosh qizi**

*3<sup>rd</sup> year student(FTO)*

*Renaissance Educational University*

*eshmatovaumida01@gmail.com*

**Abdurakhimova Nigora Bakhtiyarovna**

*ESL Teacher*

*Renaissance Educational University*

*n\_abdiraximova@renessans-edu.uz*

**Abstract:** *This paper examines the pedagogical potential of podcasts and video blogs (vlogs) as creative digital environments for developing speech skills among English language learners. Drawing on theories of multimodal literacy, authentic input, and creative language production, the study investigates how engagement with podcasts as a listening resource and vlogs as both a receptive and productive medium supports the development of pronunciation, fluency, lexical range, and discourse competence. The paper further explores the role of learner-generated podcast and vlog content as a form of creative output that integrates speaking, planning, and self-monitoring skills. Evidence from recent empirical studies suggests that systematic integration of podcast and vlog activities into EFL instruction produces measurable gains in oral fluency, prosodic awareness, and communicative confidence, particularly when combined with structured reflection and peer feedback.*

**Keywords:** *podcasts, vlogs, speech skills, EFL, multimodal literacy, authentic input, oral fluency, creative production, digital storytelling, prosodic awareness*

**Аннотация:** *В данной статье исследуется педагогический потенциал подкастов и видеоблогов (влогов) как креативной цифровой среды для развития речевых навыков у изучающих английский язык. Опираясь на теории мультимодальной грамотности, аутентичного ввода и творческой языковой продукции, исследование изучает, каким образом работа с подкастами как ресурсом для слушания и с влогами как рецептивным и продуктивным медиа способствует развитию произношения, беглости речи, лексического диапазона и дискурсивной компетенции. Результаты эмпирических исследований свидетельствуют о том, что систематическое включение подкаст- и влог-деятельности в обучение EFL даёт измеримые улучшения устной беглости, просодической осведомлённости и коммуникативной уверенности.*



**Ключевые слова:** подкасты, влоги, речевые навыки, EFL, мультимодальная грамотность, аутентичный ввод, устная беглость, творческая продукция, цифровой сторителлинг, просодика

**Annotatsiya:** Ushbu maqolada ingliz tilini o'rganuvchilarda nutq ko'nikmalarini rivojlantirish uchun podkastlar va video bloglar (vloglar)ning kreativ raqamli muhit sifatidagi pedagogik imkoniyatlari o'rganiladi. Multimodal savodxonlik, autentik kirish va ijodiy til ishlab chiqarish nazariyalariga tayanib, tadqiqot podkastlar va vloglar bilan ishlashning talaffuz, ravonlik, leksik doira va diskursiv kompetensiyani rivojlantirishga qanday ko'maklashishini tahlil qiladi. So'nggi empirik tadqiqotlar natijalari shuni ko'rsatadiki, podkast va vlog faoliyatini EFL o'qitishiga tizimli kiritish og'zaki ravonlik, prosodik xabardorlik va kommunikativ ishonchni sezilarli oshiradi.

**Kalit so'zlar:** podkastlar, vloglar, nutq ko'nikmalari, EFL, multimodal savodxonlik, autentik kirish, og'zaki ravonlik, ijodiy ishlab chiqarish, raqamli hikoya aytish

## 1. INTRODUCTION

The proliferation of digital media has fundamentally altered the landscape of language exposure available to EFL learners. Where previous generations of learners were largely confined to classroom input, today's learners have access to virtually unlimited quantities of authentic spoken English through streaming platforms, social media, and on-demand audio and video content. Podcasts and video blogs (vlogs) represent two of the most pedagogically significant of these media: they provide richly contextualised, naturally paced, register-varied speech across an extraordinary range of topics and genres, and they position learners not only as passive consumers but – through the creation of their own audio and video content – as active producers of spoken discourse.

Despite their evident pedagogical promise, podcasts and vlogs remain undertheorised and insufficiently

integrated into formal EFL instruction. Most classroom applications remain ad hoc: a teacher assigns a podcast episode for homework, or learners film a presentation without structured linguistic scaffolding or reflective follow-up. What is needed is a principled framework that identifies the specific speech skills these media develop, the conditions under which they develop most effectively, and the instructional designs that maximise their pedagogical impact. This paper contributes to that framework.

## 2. THEORETICAL FRAMEWORK

### 2.1 *Authentic Input and the Comprehensible Input Hypothesis*

The theoretical case for podcast-based instruction begins with Krashen's (1982) comprehensible input hypothesis, which proposes that language acquisition occurs when learners are exposed to input that is slightly beyond their current competence level – what Krashen designates as  $i+1$ . Podcasts, particularly



those produced for general rather than language-learning audiences, provide naturalistic input at a range of difficulty levels, allowing learners to select content that approximates their productive  $i+1$  zone. Unlike textbook recordings, which are typically scripted, simplified, and acoustically clean, authentic podcast speech contains the hesitations, reformulations, overlaps, and prosodic variation that characterise real spoken interaction – features whose exposure is essential for developing listening comprehension under genuine communicative conditions (Gilmore, 2007).

Vandergrift and Goh's (2012) metacognitive model of listening development provides a further theoretical grounding for podcast-based instruction. The model identifies planning, monitoring, evaluation, and problem-solving as core metacognitive processes in listening comprehension. Podcasts, because they can be paused, rewind, and replayed, are particularly well suited to the development of these metacognitive capacities: learners can monitor their comprehension in real time, identify points of breakdown, and deploy repair strategies without the social pressure of live interaction.

## ***2.2 Multimodal Literacy and Vlog-Based Learning***

Vlogs introduce a multimodal dimension that podcasts alone cannot provide. Jewitt's (2008) framework of multimodal literacy – which holds that meaning is made not through language

alone but through the interaction of linguistic, visual, gestural, and spatial modes – is directly relevant to vlog consumption and production. When learners watch vlogs, they process speech in conjunction with facial expression, gesture, setting, and editing choices that contextualise and enrich linguistic meaning. When they produce vlogs, they must coordinate these same modes in the construction of coherent multimodal texts, a demand that develops not only speech fluency but awareness of audience, register, and communicative purpose.

Hafner and Miller (2011), in one of the earliest empirical studies of vlog production in EFL contexts, found that learner-created digital video projects significantly enhanced discourse organisation and self-monitoring in speaking, compared to conventional presentation tasks. The researchers attributed this to the permanence of the video medium: knowing that their speech would be recorded and reviewed – by themselves, their peers, and their teacher – prompted learners to invest greater attention in planning and linguistic precision.

## ***2.3 Output, Self-Monitoring, and Creative Production***

Swain's (1995) output hypothesis provides the theoretical basis for treating podcast and vlog production as acquisition-facilitating activity. Swain argues that producing language – particularly under conditions that require precision and contextual appropriateness



– forces learners to notice gaps in their competence and to experiment with new forms. The act of recording a podcast or filming a vlog is a uniquely demanding form of pushed output: learners must sustain extended monologic or dialogic speech, manage prosody and pacing, select vocabulary consciously, and self-monitor their performance in real time. Levelt's (1989) model of speech production identifies self-monitoring as a critical mechanism through which speakers detect and repair errors – a mechanism that recording activities develop more systematically than live interaction, because the recording can be reviewed and evaluated after the fact.

### 3. PEDAGOGICAL APPLICATIONS

#### 3.1 *Podcast Listening: From Consumption to Analysis*

The pedagogical value of podcast listening is maximised when it moves beyond passive consumption to active linguistic analysis. A three-stage instructional sequence – pre-listening (activating schema and predicting content), while-listening (targeting specific phonological, lexical, or discourse features), and post-listening (reflection, discussion, and language focus) – transforms podcast episodes from entertainment into structured learning experiences. Learners can be directed to attend to specific prosodic features – stress patterns, intonation contours, rhythm – that are rarely made explicit in classroom instruction but are critical for intelligibility and listener-

friendliness in spoken English (Brazil, 1997).

Shadowing – the technique of simultaneously repeating a speaker's output while listening – has been shown to produce rapid gains in prosodic accuracy and speaking rate when applied to podcast material (Hamada, 2016). Unlike repetition drills, shadowing with authentic podcast material develops the learner's sensitivity to the full prosodic envelope of natural speech: the chunking of utterances into tone units, the placement of nuclear stress, and the management of pause and pace that mark proficient spoken English.

#### 3.2 *Vlog Viewing: Discourse Modelling and Cultural Awareness*

Vlog consumption offers specific pedagogical affordances not available in audio-only podcasts. The visual channel provides paralinguistic cues – gesture, gaze, facial expression, proxemics – that contextualise spoken meaning and model the embodied dimensions of communicative competence. Learners who regularly watch vlogs produced by native or proficient speakers develop an implicit understanding of how speech is coordinated with body language, how topics are introduced and closed, and how speakers manage the discourse structure of extended monologue – all dimensions of communicative competence that are difficult to teach explicitly but can be absorbed through extensive multimodal exposure (Jewitt, 2008).

#### 3.3 *Learner Podcast and Vlog Production*



The most pedagogically ambitious application of these media is learner-generated content: having students produce their own podcasts or vlogs as extended speaking tasks. The production process involves multiple speech skill components: planning and organising content (discourse competence), selecting appropriate vocabulary and grammar (grammatical competence), calibrating register to audience (sociolinguistic competence), and managing breakdown and repair (strategic competence). Cope and Kalantzis (2009), in their framework of multiliteracies, argue that creative multimodal production of this kind develops not only technical skills but critical awareness of how meaning is constructed and communicated – a meta-communicative understanding that enriches all subsequent language use.

Practically, learner podcast and vlog production is most effective when embedded in a structured cycle: topic selection and research, scripting or planning (without full scripting, to avoid reading aloud), recording, self-review using a structured checklist, peer feedback, and final publication or sharing. The self-review phase is particularly critical: research by Tian and Wang (2010) demonstrates that learners who review their own recorded speech and identify specific prosodic and fluency targets outperform those who produce recordings without structured reflection, both in subsequent recording tasks and in live speaking assessments.

#### 4. EMPIRICAL EVIDENCE

The empirical literature on podcast and vlog use in EFL instruction, while relatively recent, is consistently encouraging. Ducate and Lomicka (2009), in a landmark study of learner-produced podcasts in university French and German courses, found significant improvements in pronunciation accuracy and speaking confidence over a semester, with learners reporting that the recording process made them more conscious of their speech production than any other classroom activity. Importantly, the gains were strongest for learners who engaged in multiple recording cycles with feedback, suggesting that iteration and reflection are essential mediating variables.

With respect to vlogs specifically, Gromik (2012) conducted a controlled study in a Japanese university EFL context, finding that students who produced regular video recordings of their speaking demonstrated significantly greater gains in speaking length, lexical diversity, and self-correction frequency compared to a control group that practised speaking through conventional pair and group tasks. The vlog group also reported lower speaking anxiety in post-treatment assessments, a finding consistent with the protective function of the camera-mediated interaction: the absence of a live interlocutor reduced the immediacy of social judgement while preserving the communicative stakes of the task.

More recent research has examined the integration of podcast



consumption and production within a single instructional sequence. Chan, Chi, Chin, and Lin (2011) found that learners who both listened to model podcasts and produced their own showed greater metalinguistic awareness – an ability to reflect analytically on their own and others' speech – than those who engaged in either activity alone. This finding supports the theoretical prediction of Swain's output hypothesis: the combination of comprehensible input and pushed output is more acquisition-facilitating than either in isolation.

## 5. CONCLUSION

Podcasts and vlogs constitute a genuinely creative environment for speech skill development precisely because they position learners simultaneously as consumers and producers of authentic discourse. As consumers, learners gain access to the prosodic, lexical, and discursive richness of natural spoken English across an unlimited range of topics and registers. As producers, they engage in the full

cycle of speech production – planning, encoding, monitoring, and repairing – under conditions that combine communicative authenticity with the psychological safety of the recorded medium.

For EFL practitioners, the pedagogical implications are clear. Podcast listening should be structured around specific linguistic targets – prosodic features, discourse markers, register variation – rather than deployed as unguided homework. Vlog viewing should be accompanied by analysis of multimodal meaning-making, not only comprehension checking. And learner content production should be embedded in iterative cycles of planning, recording, review, and feedback that develop the self-monitoring capacities at the heart of long-term speech skill growth. When these conditions are met, podcasts and vlogs are not merely supplementary resources but transformative pedagogical tools.

## REFERENCES:

- Brazil, D. (1997). *The communicative value of intonation in English*. Cambridge University Press.
- Chan, W. M., Chi, S. W., Chin, K. N., & Lin, C. Y. (2011). Students' perceptions of and attitudes towards podcast-based learning: A comparison of two language podcast projects. *Electronic Journal of Foreign Language Teaching*, 8(1), 312-335.
- Cope, B., & Kalantzis, M. (2009). "Multiliteracies": New literacies, new learning. *Pedagogies: An International Journal*, 4(3), 164-195.
- Ducate, L., & Lomicka, L. (2009). Podcasting: An effective tool for honing language students' pronunciation? *Language Learning and Technology*, 13(3), 66-86.
- Gilmore, A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching*, 40(2), 97-118.



- Gromik, N. A. (2012). Cell phone video recording feature as a language learning tool: A case study. *Computers and Education*, 58(1), 223-230.
- Hafner, C. A., & Miller, L. (2011). Fostering learner autonomy in English for science: A collaborative digital video project in a technological learning environment. *Language Learning and Technology*, 15(3), 68-86.
- Hamada, Y. (2016). Shadowing: Who does it work for? *Language Teaching Research*, 20(2), 166-191.
- Jewitt, C. (2008). Multimodality and literacy in school classrooms. *Review of Research in Education*, 32(1), 241-267.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (pp. 125-144). Oxford University Press.
- Tian, J., & Wang, Y. (2010). Taking risks and making mistakes in speaking: An innovative approach through podcasting. *International Journal of Technology in Teaching and Learning*, 6(2), 113-130.
- Vandergrift, L., & Goh, C. C. M. (2012). *Teaching and learning second language listening: Metacognition in action*. Routledge.