

# A SECOND LANGUAGE ACTOR'S SCHIZO-AFFECT TRAINING USING AI: *THE HUMAN VOICE*

Bogdan Florea

DOI: <https://doi.org/10.14712/2571452X.2025.70.10>

**Abstract:** This article explores the use of artificial intelligence (AI) in training second language actors through the lens of schizo-affect, a concept derived from Gilles Deleuze and Félix Guattari, representing the non-linear, affect-charged dimensions of utterance. Conventional voice and speech training prioritizes native-standard fluency, often stifling individuality and the expressive potential of second language actors. Drawing on Jean Cocteau's monodrama *The Human Voice*, the project investigates how second language actors can develop a distinct, performative English – an English-becoming-another-language – by harnessing schizo-affect. AI tools such as Speakometer, Speechling, and Resemble AI, typically designed to normalize pronunciation, are repurposed to amplify the expressive, affective dimensions of speech, providing critical feedback, mirroring, and contrast. Through targeted engagement with AI, actors explore internal auditory models, isolate and manipulate idiosyncratic pronunciation, and cultivate a unique vocal identity, while simultaneously aligning schizo-affect with character creation. The analysis demonstrates that second language expression should not be treated as a deficit but as a source of creative rupture, enabling actors to integrate unpredictability, bodily tension, and unusual rhythms into performance. Ultimately, the article argues for a rethinking of second language acting pedagogy, emphasizing expressive experimentation, individuality, and the cultivation of non-standard voices as aesthetic and performative resources.

**Keywords:** actor training, artificial intelligence, second language expression, schizo-affect

## Introduction

In most conventional paradigms of second-language English usage, particularly within the domain of acting and performance, fluency is often equated with authority and correctness. Accent, intonation, and grammatical consistency are measured against standardized models, which establish themselves through

certain traditions and styles (for example, those of acting Shakespeare) or through specialized teaching, such as voice and speech training in drama schools and acting conservatoires, in which certain “correct” ways of pronunciation and delivery are taught. Deviations from these norms – speaking/performing with a “non-native” accent being one of them – are frequently understood as deficiencies that need to be corrected. For actors who use English as a second language, this model imposes a linguistic pressure that is technical, political and creative: the demand to sound “native” seems to be not just about clarity, but about legibility and legitimacy within a system that privileges certain voices over others.

Ileana Gherghina and Bogdan Florea, two UK-based Romanian-born theatre-makers who created Nu Nu Theatre in 2012, advocate for a better representation of second language actors in the industry. They support a reconceptualization of training/acting in a second language as a more fluid and affect-imbued process, which does not have to conform to a given standard. They favour the idea of decentralizing voice and speech training to recognize uniqueness and individuality. Training towards sounding “native” is seen as counter-productive and limiting for second language actors. As such, one of the aims is to articulate alternative ways of speech training and performing in the second language English. The project based on Jean Cocteau’s text *The Human Voice* represents a step in that direction, exploring (self-)correctionist pressures tilted toward “perfect” fluency in the English language. To aid that exploration, the makers introduced the Deleuze and Guattari-inspired concept of schizo-affect, which signifies the non-linear, affect-charged aspect of the utterance in the second language, which inherently resists standardization and contains creative capital. This article reflects on how AI tools have been employed to train the actors’ schizo-affect-ive, non-standard English but also in the character creation process. The makers wanted to test how AI tools, whilst normative by design, can be hacked, subverted, and tactically redeployed in service of revealing schizo-affect.

Artificial intelligence – specifically, speech-training platforms like Speakometer, Speechling, or Resemble AI – becomes a paradoxical training tool: while encouraging standardization and teaching the “non-native” to speak more “native,” it can be tactically repurposed to trace and amplify the expressive intensity of schizo-affect-ive second language use. The goal of *The Human Voice* explorations is not to discard AI as a potential training tool, but to employ it differently and feed that experimentation into the staging of Jean Cocteau’s text. By engaging with AI tools critically, second language actors can create training feedback loops that discourage homogenization and instead cultivate idiolect, glitch, speech diversity and multiplicity, personal and distinct voices in the second language.

## **Affect and Schizo-affect**

In its Deleuzian definition, affect is the change or intensity generated when bodies are modified from without or within: "To the relations composing, decomposing or modifying an individual, there corresponds intensities that affect it, augmenting or diminishing its power to act. These intensities come from external parts or from the individual's own parts."<sup>1</sup> Affects are transitional states or modes representing vital forces by which the body strives against other forces which act on it and continually resist it or hold it in check. Affects manifest as decreases and increases, augmentations or diminutions of the body's lived power, of its vital force.

Brian Massumi, the English translator of Gilles Deleuze and Félix Guattari's *A Thousand Plateaus: Capitalism and Schizophrenia*, clarifies further. Differentiating between affect and affection, Massumi explains that neither word denotes a personal feeling:

*L'affect* is the ability to affect and be affected. It is a pre-personal intensity corresponding to the passage from one experiential state of the body to another and implying an augmentation or diminution in that body's capacity to act. *L'affection* (Spinoza's *affectio*) is every such state considered as an encounter between the affected body and a second, affecting body, with body taken in its broadest possible sense to include "mental" or ideal bodies.<sup>2</sup>

To this, we can add corporeal, spiritual, animal, mineral, vegetable, conceptual, organic or inorganic, mental, ideal, physical, aural, and in the context of our investigation, linguistic bodies.

Situating affect in the realm of language and speech, we can find diverse examples: for instance, lullabies or little rhymes that may affect the body of the baby by enveloping it and producing drowsiness, eventually inducing sleep. Likewise, certain words or formulas that will instead instill a sense of discipline in a child or a soldier or an employee, illustrating how utterance, speech, words, language can act upon bodies, affecting their ability to act, diminishing or disciplining their lived power. Similarly, words, tones, expressions that can induce a state of inebriation or sexual arousal; chants, mantras, incantations, or prayers that affect individuals or crowds; speech and words that abuse,

<sup>1</sup> Gilles Deleuze and Félix Guattari, *A Thousand Plateaus: Capitalism and Schizophrenia*, trans. Brian Massumi (Minneapolis, MN: University of Minnesota Press, 1987), 256.

<sup>2</sup> Brian Massumi, "Notes on the Translation and Acknowledgements," in Deleuze and Guattari, *A Thousand Plateaus*, xvi.

diminishing that person's powers to (re)act freely. Therefore, in the context of speech and language, we consider affect to be that intensity released or created in the body of the utterer or another receiving body by a body of language. Affect is distinct from feeling, being "an active discharge of emotion, the counterattack, whereas feeling is an always displaced, retarded, resisting emotion."<sup>3</sup> In other words, affect is the cost exacted on our bodies with almost every utterance that we produce or receive. It is an electric discharge in the body: a pre-personal intensity that has a distinct performative dimension.

Within the wider realm of language-induced affect, the focus is on the particular type of affect generated when a body utters not in the mother tongue but in the second language, which we call schizo-affect. When trying to define schizo-affect a rather extreme example comes to mind: a cat, a husky dog or parrot trying to produce human language – the internet abounds in short videos showing such examples. It is useful to observe the non-human animals' strain and effort to produce comical, semi-accurate versions of an "I love you," swear words or sometimes trying to sing. These non-human animals are engaged in a schizo-utterance that brings together – requiring tremendous effort and poise – two bodies: that of the non-human animal, with its ability to emit certain sounds and that of the human language. Such an undertaking releases in the body of the non-human animal a performative intensity: a tearing asunder becomes visible in how that utterance is physicalized, in the distorting of the body to obtain the sounds, etc. The utterance consumes and is costly to the non-human animal, triggering a specific affective jolt in the listener. To a lesser degree, the second language speaker/actor embodies a similar schizzed affect.

Therefore, we observe our and others' occasional slip-ups, decelerations, over-expansive or on the contrary collapsed gesticulation, a- or anti-rhythmicities, hiatuses, stutters, breathlessness, disharmonious or, on the contrary, bodily transformations reflecting the uneven juxtaposition of the bodies of the first and second language within the performative space of a single utterance. Examples of such second language performative acts: a driver informs passengers on a London Victoria-bound coach about safety regulations and rules on board. He constructs his micro-performance on the microphone almost as a performance piece: well-rehearsed over so many journeys yet still somewhat precarious, outside his comfort zone, delivered in a theatrical tone, trying to avoid or cover jerky, shaky enunciation. Similarly, a *Big Issue* seller, with a wailing invitation to passers-by to buy her magazine, her speech in a broken English gaining mechanical, almost incantatory dimensions. Likewise, the sales assistant in a shop or over the phone,

<sup>3</sup> Deleuze and Guattari, *A Thousand Plateaus*, 400.

tripping over words and making grammatical errors in English. Such second language acts or micro-mis-en-scènes fascinate through unusual bodily postures, fretting, theatrical gesticulation, exaggerations that seek to compensate for the lack of native-like pronunciation or syntax, hiatuses. These performances are based not solely on feeling, emotion, thought, competency, social status, empathy, communication of facts but also contain the tearing-asunder intensity of schizo-affect. Why schizo? Because it is created at the intersection of two different language systems – the mother tongue and the second language – within the performative space of embodied utterance.

The schizo-affect materializes as a sonorous and bodily choreography/performativity of intensities, tensions, cacophonies, unexpected vaults or sudden craterings in a body's power to act. Schizo-affect appears as an unavoidable tension between the second language and the habitus of the native language, which cannot be fully abandoned and thus leaks and seeps in utterance into the second language: "Language is a body technique, and specifically linguistic, especially phonetic, competence is a dimension of bodily hexis [...]"<sup>4</sup> Whilst the first language is fully inscribed in the hexis, the second language appears as a costume, a second skin, as it were: something that the body needs to adjust to. Associating two (sometimes very different) phonetic and prosodic systems creates tensed discrepancies and hybridizations, which have a distinct performative dimension.

Schizo-affect is the pulling asunder that characterizes how speakers perform (themselves) socially, culturally, artistically, etc., in a second language. Such performing is not realized solely by means of controlling and communicating feeling or emotion, accurately conveying facts or information, employing conventional body language, but also by trying to master the fleeting unknown of schizo-affect, which unfolds in speech as a sum of unpredictable intensities and reactions that result from an imperfect, fluid suture, which conjoins two distinct blocks of pronunciation, prosody, etc. The intensity or tension released invariably leaks – often unexpectedly or unconsciously – into the utterance, surfacing as glitch, eccentricity, augmentation or diminution. Influencing behaviour and embodiment, schizo-affect is thus fundamentally performative and often distinctively theatrical.

### **Schizo-affect, Not Accent**

Introducing the term schizo-affect helps to intensely nuance the idea of "non-native" accent. Accent appears as a totalizing, centralizing concept which brings

<sup>4</sup> Pierre Bourdieu, *Language and Symbolic Power*, ed. John B. Thompson, trans. Gino Raymond and Matthew Adamson (Cambridge: Polity Press, 1991), 86.

specific sonorities or pronunciation traits under a very big umbrella. Categories such as French accent, Eastern-European or, worse, Indian or Chinese accent risk to erase entire linguistic and cultural ecosystems. They invite stereotyping (especially in acting and theatre-making), by ignoring speech individuality and uniqueness. As actors, we instinctively reject the description of characters through wide categories such as accent, arguing that actors/acting can never be reduced to playing just (their) accents: first and foremost, actors perform characters, human lives, and destinies. The introduction of schizo-affect seeks to address those risks. With schizo-affect, we can concentrate on the very personal and poetic manner in which a certain individual/actor/character (not a whole group or a category) speaks/performs. Each character played demands a personal, unique, distinct and individual patois, a targeted version of English, an English-becoming-another-language or a schizo-affective English languaging, as it were. This patois or personal, unique micro-accent accent is determined both by the actor's language skills in general and by the role played. Schizo-affect shifts the focus thus on a creative resistance against native standardization and against hierarchical, hegemonic accentism, by accepting the apparent deviations – hesitations, tonal asymmetries, physical/bodily alterations and incongruences – not as limitations, but as aesthetic resources and poetic dimensions of one's identity in the second language, transferrable to the characters created. Schizo-affect means that there should be no process of repair or remedy one's "non-native" accent, but a practice of rupture, deterritorialization and schizzing, constructing thus the actor's unique mark in the second language (like Garbo's English, Banderas's English, etc.). Training the schizo-affect implies being playful and manipulative with the language, either in phonetics, syntax, or pragmatics<sup>5</sup> with a view to discovering new expressive territory, limited otherwise by the fixity of the "native" standard of speaking.

### ***The Human Voice***

Our training experiment was combined with work on staging Jean Cocteau's *The Human Voice*. First produced by the *Comédie-Française* in 1930, the monodrama is set in Paris, where a still-quite-young woman is on the phone with her lover of the last five years, who is breaking up with her. He is to marry another woman the next day, which causes the protagonist to despair. The monologue triggers the woman's crippling depression, which unfolds step-by-step over the phone

<sup>5</sup> Nancy Bell, "Formulaic Language, Creativity, and Language Play in a Second Language," *Annual Review of Applied Linguistics* 32 (2012): 189-205.

towards a heart-wrenching climax. For the duration of the performance, the audience sees and hears the woman on stage speaking on the telephone with her invisible and inaudible departing lover, as the sense of isolation, confusion and helplessness grows. The audience can only deduct or imagine what the man at the other end of the line is saying, as he never appears on stage. The first thing that attracted us to Cocteau's text was the way in which loneliness, abandonment, the sense of deep solitude, sadness, and desperation emanate powerfully and somehow naturally from the protagonist's long monologue.

A second theme of particular interest was that of technology, as the play also invites thinking about how technology may aid or disrupt our most intimate human relationships. The frequent interruptions, cut-offs, glitches that occur during the woman's telephone conversation support the idea that technology cannot replace unmediated, uninterrupted, fluid physical contact and dialogue. The woman's agonizing dependency on the device; the vital exchanges interrupted or derailed by the caprices of the telephony signal are stark reminders of how we are in desperate need for human contact and closeness, and that no other medium can substitute that. In this play, we see the machine, and by extension technology, functioning both as a tormenting instrument (the woman is being left hanging on the phone, in a cold, impersonal, indirect manner; she is being lied to; interrupted, etc.), as a screen behind which to hide (the man perhaps does not have the courage to discuss the break up in person) and also, somewhat paradoxically, as an instrument that keeps the woman alive (the man's voice at the other end, as long as it keeps talking, is a source of life): "despite everything we are attached to each other by the telephone. [...] I was connected to you by a diver's air tube [...]. Now I have air to breathe because you are talking to me."<sup>6</sup> The text, seen through a twenty-first century lens, reminds us how technology and machines have changed and will continue to affect human relationships. Responding to these two themes in particular, we saw Cocteau's text as lending itself very well to our schizo-affect explorations using AI.

Keeping schizo-affect in focus, our question is if we were to use AI tools such as Speakometer or Speechling (instead of a human accent/speech coach) to help us work with English for this play, what would we be able to discover? How would training with AI help us better preserve and frame the uniqueness, distinctiveness and individuality of our English-becoming-another-language? Can AI process schizo-affect? Can working with AI help us find our voice?

<sup>6</sup> Jean Cocteau, *The Human Voice* (La voix humaine, 1930), trans. Carl Wildman (London: Vision Press, 1951), 10-11.

The notion of (the human) voice functions as “a multifaceted metaphor that encompasses the ideas of agency, identity, authorship, ownership, writing [and speaking] style, linguistic register, rhetorical stance and textually embodied self.”<sup>7</sup> Using AI as a modality of analysis, expression, construction and bringing forth our voice draws on a text about how someone’s voice can be a source of life, a sharp, incisive tool, the bridge that connects two people, two lives, an instrument with which one starts their journey into death. In the stage directions, Cocteau indicates that: “*The actress should give the impression that she is bleeding, losing her life’s blood, like a wounded beast, finishing the play as if the bedroom is drenched in blood.*”<sup>8</sup> The instruction here is that the actor, through the means of voice and words, should, as the play progresses, depict the slow, agonizing bleeding towards a state of death or near-death. Bleeding through words and with the voice is, in a sense, the main acting task that the actor must gradate across her performance. This is an immediate link that can be created between Cocteau’s protagonist’s language and voice performance, and our language and voice training as second language actors.

In similar fashion to Cocteau’s character, we too sought to realize a crescendo in how we devised our AI training: firstly, working with AI on what we felt as trickier to pronounce words and sounds, moving then on to sentences in which we played with pronunciation, rhythm and pitch. Thirdly, we experimented with cloning an AI voice that would be able to replicate the actor’s idiosyncrasies, tonalities, and non-standard pronunciation. During the process, we noticed that AI could be used as a mirror and for contrast.

### **AI: Mirror and Contrast**

AI can offer a hyper-normative reflection, a kind of vocal mirror. This may help isolating what the actor thinks schizo-affect *is not*, sharpening an awareness of what makes their speech unique. In working with AI for *The Human Voice* project, we felt that we could train to become better at identifying, challenging and deconstructing normative pressures and expectations, illuminating the schizo-affect as an important performative capital buried in utterance, which once saved and controlled can be channelled into the character itself.

Conventional AI tools like Speakometer, Speechling, Resemble AI are typically built around a normative model of “correct” English pronunciation and fluency. These platforms use large datasets of standardized (often General American or RP)

<sup>7</sup> David I. Hanauer, “Measuring Voice in Poetry Written by Second Language Learners,” *Written Communication* 32, no. 1 (2014): 69.

<sup>8</sup> Cocteau, *The Human Voice*, 2.

speech to evaluate users and recommend corrections. Their feedback is quantitative, emphasizing fluency, intonation, pronunciation accuracy, and sometimes prosody – but always in reference to the established norm. When the actor feeds their schizo-affect-ive English into these tools, it should be expected that they “flag” expressive divergences as errors or imperfections, making the actor more immediately aware and easier for them to incorporate into speech and performance.

With AI acting as a powerful, depersonalised embodiment of the existent normative gaze, the actors’ job is to recognize the standardizing tendencies and develop skills to creatively (sometimes sarcastically) counter react, coming up with circumventing strategies. The actor must learn how to intentionally preserve, shape and play with “erroneous” parts of the speech, absorbing these thereafter into the structure of the character. AI tools, although they may subtly encourage homogenization and centralization around a standard, risking flattening the actor’s idiosyncratic expressivity, can also act as a useful mirror, inciting the actor to give contour to those idiosyncrasies. Not attuned to poetic, aesthetic, or schizo-affect-ive dimensions of language that fall outside fluency metrics, AI tools can be treated as cold hard surfaces that the actor’s speech hits, offering an effective contrast surface onto which to bring into relief schizo-affect. The fact that AI tools have difficulties detecting small pronunciation differences makes the actor’s critical and tactical work easier and more relaxed.

In our work, we have used mostly Speakometer which accommodated vocal mirroring and self-imitation. It enabled us to access our internal auditory models in English: “Sensorimotor learning theory [...] posits that speech development is guided by internal auditory models used to adjust output over time.”<sup>9</sup> In other words, we were able to learn to speak the second language by fine-tuning our non-standard internal voices (how one hears oneself speaking the second language inside). We could listen better and more intently, attuning internally our imperfect, non-normative voice rather than imitating the native standard coming from without. During work, the focus had to shift to the internal auditory second language models, and we imagined ourselves as tennis players who train their reflexes and skills, by repetitively hitting the ball against a wall. In the process, we created a mechanics that bounced us back to the internal auditory voice model. By playing against the wall, as it were, the player understands themselves better as a player. By bouncing her utterance against the AI wall, the actor can hear themselves better in the second language. Engaging in a routine of self-imitation

<sup>9</sup> Ewa Kusz and Judyta Pawliszko, “Efficacy of Artificial Intelligence (AI) Voice Cloning in Phonetic Self-Imitation for L2 Pronunciation Training,” *International Journal of Applied Linguistics* (2025): 3, <https://doi.org/10.1111/ijal.70014>.

of this nature we thought aided us in developing the internal personal voice, which corresponds with the claim that

self-imitation based on the acoustic adjustment of the learners' speech input reflects native-like pronunciation features while retaining their unique vocal identity. This personalized feedback is more effective than direct imitation of native speakers, as it increases familiarity and prompts more natural and sustainable learning.<sup>10</sup>

With both speech training and character creation, we felt that working with the non-human AI (compared to a human coach, director or collaborator) allowed for a judgement-free and lower pressure work process. We became more familiar with our second language voice, whilst AI kept functioning as a benign assistant, not inducing anxiety or hesitation. Always constant, the AI gave sufficient space to hear ourselves. Distinct from a human coach, with whom it is more difficult to achieve a non-judgmental and non-pressurized interaction, AI allows the actor to have more freedom of exploration and schizo-affect-ive improvisation in the English language, perhaps with increased enjoyment, playfulness and willingness.

The actor playing The Woman character has used AI to work with sounds and words from the text that felt trickier to utter. For instance, the word "robot" from the line "I get dressed, I go out. I come back, like a robot,"<sup>11</sup> presented some difficulty. Reflecting on the ping-pong with AI, the actor commented on how she felt that her diaphragm, jaw, tongue, larynx, all carried habits of force, not just positions of articulation. When the brain is busy making meaning, your mouth reverts to what it knows best, to its instinctive choreographies. Whilst the interaction with the AI may help smooth over some of the pronunciation difficulties, the heavy bodily tension triggered by a word such as "robot" cannot be fully solved. Schizo-affect appears when there are two unequal language bodies vying to get embodied simultaneously. The actor thought that the force thus released could not be eliminated/overcome. During the play, Cocteau's protagonist uses words to cover her desolate state and what she really feels, and the actor's task is to render this added tension. The solution was to employ some of the heaviness in uttering, the elongated flow of the text during trickier parts as a way to signal the fact that the character is lying or covering up things that she finds very difficult to say directly. Schizo-affect metamorphosed into the difficulty with which the character communicates to her lover.

<sup>10</sup> Kusz and Pawliszko, "Efficacy of Artificial Intelligence (AI) Voice Cloning," 2.

<sup>11</sup> Cocteau, *The Human Voice*, 3.

Also, the actor observed that when you repeat or work on a single word, which appears to the actor like an individual vowel and consonant cluster/entity, you're treating it almost as a musical note, which you try reaching over and over again. Working on an individual sound or word, the actor can more easily control articulation and the uttering apparatus. Working on a single word was somehow mechanical and therefore the actor felt she could "nail it." The actor discovered that the artificially rehearsed articulation and tone, could be grafted onto the character's way of speaking, helping depict the false way in which she affirms that her life is absolutely normal despite the break-up and deep loneliness. It is significant to note that the actor shaped her internal voice/auditory models in relation to the text studied. The AI training was not conducted in general but targeted towards creating an English-becoming-another-language, or a schizo-affect-ive English for the woman in *The Human Voice*. Schizo-affect was harnessed so that it rhymed with the needs/traits of the character, whilst fitting more generally with the themes and performative undercurrents of Cocteau's text. In that respect, working with an AI assistant seemed to speak directly to one of the main themes of the text: loneliness and isolation. The protagonist is alone on the stage speaking on/into a telephone, to a person with which she is not in physical contact or proximity. The actor also trained/performed alone with an AI, paralleling the character's condition. More specifically, the actor does not rehearse her voice in English in general but in connection (this time) to the words of a lonely, despairing woman.

As the training and staging experimentation progressed, more advanced tools were used, supporting the actor to train her own voice model. This expanded the possibility of creating a "self-reflective" AI, with a stronger and more wholesome mirroring capacity. At this stage, the AI tool can begin to learn and reproduce the actor's rhythms and affective gestures, allowing for more experimental uses. Tools such as Praat, an Open-source Phonetics Software, can be used to visualize pitch, formants, breath breaks, speech rhythms. We have chosen to start our experimentation with Resemble AI, which is a tool for training an AI voice based on a person's own speech. The actor used it to train a model of her schizo-affective English with the view to then be able to hear herself speak through an AI double: a de-personalized/ghosted version of the actor's voice. The model voice thus created allowed the actor to follow more easily and carefully her speech's waveforms and attempt to distance herself from her voice (hearing herself from the outside as it were). Obtaining such detachment, the actor could follow her speech's waveforms as an audience member, concomitantly, mentally and physically shaping the wider schizo-affective map of her role. Based on the AI rendition of the voice, the actor can devise or direct/stage a personal schizo-affect choreography of their role.

For example, the actor worked on the following passage:

Oh! my darling, don't apologize, it's quite natural, I'm the one who's being stupid. ... [in this sentence, the voice of the actor was almost inaudible, and the words became barely comprehensible] You are sweet. [...] Neither did I, I never thought I could be so strong. ... You mustn't admire me. I go around a bit like a sleepwalker. I get dressed, I go out. I come back, like a robot. Perhaps I'll be less brave tomorrow. ... You?<sup>12</sup>

AI synthesized the voice based on the above passage and others, and the actor was able to play back and listen to it detachedly and critically. The actor observed how the native language seemed to want to reassert itself, not because the utterance in the second language had failed but because the body seems to remember through rhythms and not through pronunciation rules. It cannot completely forget its original speeds and intensities. The AI voice was sufficiently successful in replicating the rhythm and intonation of the actor, having satisfactorily recorded and integrated some of the pauses, hesitations, whispers or non-audible passages. However, the actor did feel a considerable sense of estrangement from the synthetic voice, whilst noting that she felt that the voice's English sounded much better than how she was hearing herself inside.

Once the model voice started to take shape, the actor benefitted more from the critical distance from their own utterance and could start concentrating on constructing a thread of schizo-affect flowing through the entire play. The training/performing technique introduced a machinic double of the actor's voice in the second language, allowing the actor time to pause and employ schizo-affect in the same way as movement, physical attitude or gestures: as expressive means for portraying the character. Whispering or half-muting the text became strategies by which to overcome difficulty of pronouncing certain successions of sounds. The physical difficulty, discomfort or frustration generated; the bodily awkwardness of not being able to speak those particular sounds/words/successions was distilled into gesture and physical movement (the character is described as frantically playing with the telephone cord or shifting positions in bed). When merged with schizo-affect generated through speech, such choreography injects credibility into the character. Slowing down, taking deeper breaths, pausing, expanding the body to rhyme with the difficulty in pronunciation, twisting or shrinking the body to respond to alterations in rhythm and pitch also become justified as means of portraying a woman who is slowly dying in front of the audience.

<sup>12</sup> Cocteau, *The Human Voice*, 3.

## **Conclusion**

The initial questions were focused on how AI can help better preserve and frame the uniqueness, distinctiveness and individuality of English-becoming-another-language, on the way AI can be compared to a human coach, how does it process schizo-affect, and if working with AI helps with finding one's voice.

The experimentation with AI tools proved our central argument: that second-language expression is not a problem to be fixed, but a source of creative rupture. To discover their voice in the second language, actors should embrace schizo-affect in all its fluidity and unpredictability and allow it to work on their bodies and concomitantly shaping the character. The fluidity and unpredictability can be curated into schizo-affect-ive maps that branch out across the entire performance text, enriching the character played.

AI, when repurposed critically, can become a good ally in cultivating schizo-affect-ive speech and highlight individuality and uniqueness of the actor's voice in the second language. AI can complement human training (acting/voice/speech coaches) by providing a cold mirror in which the voice of the actor reflects itself, becoming ever more present. Through the creation of an AI-generated ghost voice, the actor can gain increased critical distance and detachment from her speech, which, in turn, helps with integrating the non-conforming voice into the tissue of the character portrayed.

Following this type of experimentation, we are inclined to rethink second-language performance and training – not as an exercise of correcting a deficit, but as tapping into an expressive reservoir that resists standardization and opens new expressive and creative possibilities. Our research also shapes itself into a call to action: to build new tools, methods, and pedagogies that honor linguistic difference and aesthetic/poetic experimentation. These pedagogies can contribute to achieving novel approaches and understandings of texts and characters, leading towards a more comprehensive poetics of foreignness and difference.

## **BIBLIOGRAPHY**

- Bell, Nancy. "Formulaic Language, Creativity, and Language Play in a Second Language." *Annual Review of Applied Linguistics* 32 (2012): 189-205.
- Bourdieu, Pierre. *Language and Symbolic Power*. Edited by John B. Thompson. Translated by Gino Raymond and Matthew Adamson. Cambridge: Polity Press, 1991.

- Cocteau, Jean. *The Human Voice*. Translated by Carl Wildman. London: Vision Press, 1951.
- Deleuze, Gilles and Félix Guattari. *A Thousand Plateaus: Capitalism and Schizophrenia*. Translated by Brian Massumi. Minneapolis, MN: University of Minnesota Press, 1987.
- Hanauer, David I. "Measuring Voice in Poetry Written by Second Language Learners," *Written Communication* 32, no. 1 (2014): 66-86.
- Kusz, Eva and Judyta Pawliszko. "Efficacy of Artificial Intelligence (AI) Voice Cloning in Phonetic Self-Imitation for L2 Pronunciation Training." *International Journal of Applied Linguistics* (2025): 1-18. Accessed 20 October 2025. <https://doi.org/10.1111/ijal.70014>.
- Massumi, Brian. "Notes on the Translation and Acknowledgments." In Gilles Deleuze and Félix Guattari, *A Thousand Plateaus: Capitalism and Schizophrenia*, xvi-xix. Minneapolis, MN: University of Minnesota Press, 1987.