From Overt Gesture to Covert Micro-Movement: Kinematic-Acoustic Trade-Offs and Internal Cueing in Art Song Performance

Yixiao Cao
Institute of Music, UCSI University, Kuala Lumpur 56000, Malaysia
Corresponding author: Yixiao Cao (e-mail: gaperseus@163.com)

ABSTRACT Art song performance requires singers to communicate textual and musical intensity through economical visible action. Larger rehearsal gestures may support phrase direction and breath organization, but they can appear stylistically excessive in recital contexts. This study examines the transition from overt gesture to covert micro-movement as a measurable kinematic-acoustic coordination problem. A single singer-researcher performed a phrase window from Wolf's Ganymed under three within-performer conditions: Neutral, Overt, and Compressed. Two smartphone videos were processed using OpenCap to extract left shoulder flexion, left elbow flexion, and pelvis tilt, while synchronized audio was analyzed in Praat for intensity, fundamental frequency, and harmonicity-to-noise ratio. Reflexive logs documented task focus and perceived inhibition after each take. Results show that Overt singing reorganized upper-limb behavior, especially elbow motion, while Compressed singing reduced visible arm and elbow range without simply returning to Neutral. Mean intensity and central pitch span remained broadly stable, whereas mean harmonicity decreased from 9.34 dB in Neutral to 6.38 dB in Overt and 4.95 dB in Compressed. The findings indicate that compressed gesture may preserve internal cueing but may also remove mechanical impulse or introduce inhibition load.

INDEX TERMS art song performance, kinematic analysis, vocal acoustics, internal cueing, signal processing

I. INTRODUCTION

ART song performance sits at a demanding intersection of sound, text and restrained embodiment. In recital culture, communicative force often comes from controlled stillness, small shifts of attention, and carefully scaled bodily action rather than from large theatrical display. For the singer, this creates a practical dilemma. A larger gesture may help breath organisation, phrase direction and emotional availability in rehearsal, but the same gesture may look out of proportion on a recital platform. Conversely, a visually economical performance may be vocally neutral or even costly if the singer achieves restraint by bracing.

This study uses the term economy to mean active scale control, not absence of movement. Economy asks the performer what should be externalised, what should be contained, and how much movement is enough for the task. That framing is consistent with embodied accounts of music, in which meaning is organised through action, perception and sensorimotor simulation rather than through sound alone [1,2]. It is also consistent with studies of singing posture showing that apparent stillness can contain continuous postural adjustment. Lyric singers' balance behaviour differs between quiet stance and singing tasks, and choral singers' body motion changes with ensemble task demands [3,4]. A recital singer therefore does not choose between motion and no motion. The more precise question is how motion is allocated and scaled.

The visual channel makes this question harder. Seeing a singer can support lyric comprehension, and visual information can strongly influence music-performance judgments [5,6]. Recent work also suggests that the amount of motion perceived as optimal is style-dependent; more movement is not automatically judged as better [7]. In art song, where textual intimacy and stylistic proportion matter, a singer must keep separate two kinds of effectiveness: what helps the sound and what reads plausibly to the eye. These two dimensions often overlap in performance, but they should not be assumed to be identical. A gesture may be useful as a rehearsal tool precisely because it is large enough to be felt, while the same scale may be visually intrusive in a Lied recital.

Voice and gesture research gives further reason to treat the issue as more than stagecraft. Singer-focused studies describe bodily movement as part of physical control, musical structure and expressive intention [8,9]. Biomechanical work on gesture-speech physics shows that high-impulse arm movement can interact with the respiratory-vocal system and modulate vocal intensity, fundamental frequency or timing [10,11]. Related music research shows that manual gesture and vocalisation can be coupled in performance practice, though the mechanisms and cultural functions differ by genre [12]. These findings make it plausible that rehearsal gestures can affect vocal coordination. They do not prove that a small, stage-facing version of the same gesture preserves the same effect.

Laban-based movement theory is useful here because it separates movement quality from mere size. Effort theory describes how Weight, Time, Space and Flow combine to form recognisable qualities of action [13]. Actor-training adaptations use those qualities to help performers scale, internalise or contain impulses [14]. Somatic and computational studies also show that bodily configurations can carry affective signatures [15,16]. Yet these literatures must be used cautiously in singing. They do not license a simple prescription such as adding a given Effort quality to guarantee a given sound or audience response. In art song, movement must remain compatible with breath management, laryngeal stability and stylistic restraint.

The gap addressed here is the translation from overt rehearsal gesture to compressed performance gesture. Musical-gesture scholarship treats phrases as temporally extended action-sound units rather than isolated poses [17,18]. Perception studies also show that observers attend to timing, direction, smoothness and movement character, not only to movement amount [19,20]. Historically, however, song and opera do not share the same visual regimes. The highly legible rhetorical gesture associated with staged theatre cannot simply be imported into intimate Lied performance [21]. The present study therefore asks a narrow but practical question: when a rehearsal-sized gesture is compressed into a covert or micro-movement, what changes in the body, what changes in the sound, and what kind of explanation is most plausible? This question also has pedagogical consequences. Singers are often encouraged to use generous arm paths, torso release or imagined spatial trajectories during technical work. Later, they are asked to make the result stage-appropriate. The missing step is an explicit way to examine whether the compressed form still carries a useful cue or whether it becomes a new source of effort. Without that step, gesture can easily be discussed in vague terms: natural, too much, restrained, expressive or distracting. The present article converts that vague discussion into a small measurement problem.

The article reports a single-singer, within-piece pilot comparison in Wolf's Ganymed. Three intentional conditions are compared: Neutral, Overt and Compressed. The aim is not to prove a universal effect of gesture. It is to develop a diagnostic procedure that can expose trade-offs between visible economy, kinematic organisation, acoustic descriptors and the singer's felt task.

II. EXPERIMENTAL

A. MATERIALS AND METHODS

The study used an idiographic practice-led design. The performer was the singer-researcher, and the unit of analysis was a controlled phrase window rather than a population sample. The comparison was within-performer and within-piece, so the findings are interpreted as evidence about one singer's coordination under three task constraints. They are not treated as general laws for all singers.

The selected excerpt was the 14-24 s window of a take of Wolf's Ganymed, corresponding to the phrase Es schweben die Wolken abwärts. The window was chosen because it is musically bounded, contains a clear impulse toward gesture, and was one of the most reliable synchronisation points in the available audiomotion dataset. It is long enough to include several postural and arm adjustments but short enough to remain one interpretable gestural-sonic object. This window-based approach is important because a full take can hide local coordination problems. A phrase window allows the researcher to ask what one musical moment demands and how different task instructions reshape the same moment.

Three intentional conditions were recorded. Neutral served as a baseline coordination state; it was defined not as an absence of intent, but as a state of professional vocal readiness involving standard postural adjustment without any additional, experimentally planned gestural layer. Overt enlarged the rehearsal gesture, allowing the arms and a slight forward inclination to support phrase direction. Compressed recalled the same gestural impulse while reducing its external visibility, aiming for a stage-facing micro-movement rather than a full arm action. The order was fixed as Neutral, Overt and Compressed, with a two-minute rest between takes. This order reflects a rehearsal-to-stage pathway but does not control for possible Order Effect Bias. Specifically, because the Compressed condition was consistently performed last, the observed drop in HNR may be partially confounded by vocal fatigue or the cumulative cognitive load of the preceding takes. Future iterations of this protocol should employ a counterbalanced order to isolate the gestural variable from temporal effects. Motion data were captured with OpenCap using two smartphones in frontal and lateral views. OpenCap provides video-based estimates of musculoskeletal kinematics and is suitable for accessible movement analysis outside a laboratory setting [22]. The variables analysed were left arm flexion (arm_flex_l), left elbow flexion (elbow_flex_l) and pelvis tilt (pelvis_tilt), exported as OpenSim coordinate data in degrees at approximately 60 Hz. These variables were selected to represent upper-limb organisation and coarse axial/postural configuration. They should not be read as complete descriptions of vocal technique, scapular action or breath management.

Audio was recorded with a Zoom H5 placed approximately 1 m from the performer and oriented perpendicular to the face. Praat was used to extract acoustic descriptors [23]. Intensity was computed with a 0.01 s time step and 60 Hz pitch floor. Pitch was estimated with autocorrelation, using a 60 Hz floor, 400 Hz ceiling and 0.01 s time step. Harmonicity-to-noise ratio (HNR) was extracted from cross-correlation harmonicity with a 0.01 s time step, 60 Hz minimum pitch, 0.1 silence threshold and 4.5 periods per window. Praat defaults handled unvoiced frames; unvoiced intervals were not entered as zero values.

Descriptive statistics were calculated for each condition: mean, standard deviation (SD), conventional range and robust range where relevant. Robust range was defined as q95-q05 to reduce sensitivity to brief tracking artefacts. For kinematic contrasts, moving-block bootstrap intervals were calculated with a 0.5 s block length and 2,000 resamples. These intervals describe within-take variability in an autocorrelated time series. They do not support population inference. Acoustic contrasts are reported as point estimates relative to Neutral, because the design does not contain enough independent takes to support stronger statistical claims. Immediately after each take, the singer-researcher recorded a brief reflexive log. Entries noted attentional focus, allowed versus inhibited movement, sites of effort or stiffness, and perceived vocal consequence. The log was not treated as proof of mechanism. It was used to distinguish possible explanations that cannot be separated from kinematic and acoustic traces alone, especially mechanical coupling, internal cueing and inhibition load. This is essential in practice-led voice research because the same visible action can be produced with different internal effort.

III. RESULTS AND DISCUSSION

A. EMBODIMENT, VISUAL ECONOMY AND THE REHEARSAL-STAGE PROBLEM

The literature points to a useful but unstable premise: movement can be meaningful, functional and stylistically risky at the same time. Embodied music cognition and musical-gesture research make it difficult to treat gesture as a decorative layer added after the musical work is complete. For singers, posture, breath, phrase intention and text delivery are coordinated in the same performing body. At the same time, art song performance usually narrows the acceptable visual bandwidth. The singer is visible, but not normally theatrical in the operatic sense. This is why the present data are interpreted through trade-off rather than improvement. The three conditions were designed to expose that trade-off. Neutral asks what happens when no additional gestural plan is imposed. Overt asks what happens when a rehearsal gesture is allowed to become clear and physically available. Compressed asks a harder question: what remains when the singer keeps the internal impulse but restricts the external display? The distinction matters because compression may change the function of gesture. An overt action can supply momentum, spatial direction and expressive permission. A compressed action may work more as an attentional cue. It may also become an inhibition task.

B. KINEMATIC REORGANISATION ACROSS CONDITIONS

The first result is that Compressed was not simply a smaller version of Overt. Table 1 shows that the Overt condition substantially reorganised the upper limb, especially the elbow. Mean left elbow flexion rose from 55.21 degrees in Neutral to 108.39 degrees in Overt, with the robust range also increasing from 23.53 to 52.22 degrees. The left arm variable showed a larger robust range in Overt (66.85 degrees) than in Neutral (36.74 degrees). These patterns fit the task design: Overt singing externalised phrase energy through a visible arm pathway.

TABLE 1. Descriptive kinematic variability for Ganymed in the 14-24 s phrase window

Variable	Condition	Mean	SD	Range (max-min)	Robust range (q95-q05)
arm_flex_l	Neutral	15.35	15.27	63.10	36.74
arm_flex_l	Overt	8.18	24.59	73.89	66.85
arm_flex_l	Compressed	6.79	7.94	24.49	22.09
elbow_flex_l	Neutral	55.21	6.99	59.20	23.53
elbow_flex_l	Overt	108.39	20.54	66.06	52.22
elbow_flex_l	Compressed	50.30	3.01	11.90	10.22
pelvis_tilt	Neutral	-13.35	1.14	7.06	3.70
pelvis_tilt	Overt	-15.93	1.52	7.80	4.69
pelvis_tilt	Compressed	-11.07	1.05	4.53	3.53

Compressed produced a different profile. Its left arm robust range fell to 22.09 degrees and left elbow robust range to 10.22 degrees, both lower than Neutral. Mean elbow flexion also returned close to Neutral rather than remaining near Overt. Pelvis tilt, however, did not simply return to baseline: the mean changed from -13.35 degrees in Neutral to -11.07 degrees in Compressed. This suggests a redistribution of movement, not a global deletion of movement. In practical terms, making the arms look quiet may shift the burden of organisation into axial posture or internal muscular effort.

Table 2 gives the within-take contrast intervals. The Overt-Neutral elbow contrast was large and stable: +53.18 degrees in mean elbow flexion, with a 95% moving-block interval of 44.69 to 62.18 degrees. The Compressed-Neutral elbow contrast was smaller but still negative for the mean (-4.91 degrees, interval -7.28 to -2.22 degrees). For arm flexion, Compressed also reduced mean and robust range relative to Neutral. Pelvis tilt moved in opposite directions across Overt and Compressed: Overt was more negative than Neutral, while Compressed was less negative. These values support a main kinematic finding: compression changed the coordination regime rather than merely reducing Overt by a fixed percentage.

TABLE 2. Moving-block bootstrap intervals for kinematic contrasts in the 14-24 s window

Variable	Contrast	Delta mean (deg) [95% CI]	Delta robust (q95-q05) [95% CI]
arm_flex_l	Overt - Neutral	-7.17 [-19.93, 4.42]	30.11 [11.05, 34.01]
arm_flex_l	Compressed - Neutral	-8.56 [-15.70, -1.66]	-14.64 [-32.10, -10.46]
elbow_flex_l	Overt - Neutral	53.18 [44.69, 62.18]	28.69 [20.26, 47.91]
elbow_flex_l	Compressed - Neutral	-4.91 [-7.28, -2.22]	-13.31 [-22.31, 3.45]
pelvis_tilt	Overt - Neutral	-2.57 [-3.31, -1.85]	0.99 [-0.71, 3.06]
pelvis_tilt	Compressed - Neutral	2.28 [1.65, 2.92]	-0.17 [-1.69, 0.88]

The figures visualise the same point. The shoulder and elbow traces show that Overt was not only larger but differently timed and configured. Compressed narrowed the visible upper-limb pattern, yet its curve did not become identical to Neutral. The pelvis trace confirms that even when the arms were reduced, the wider postural system still changed. This matters for pedagogy because singers often describe stage economy as if it were a simple instruction: do less. The data suggest a more useful instruction: find where the reduced gesture relocates effort.

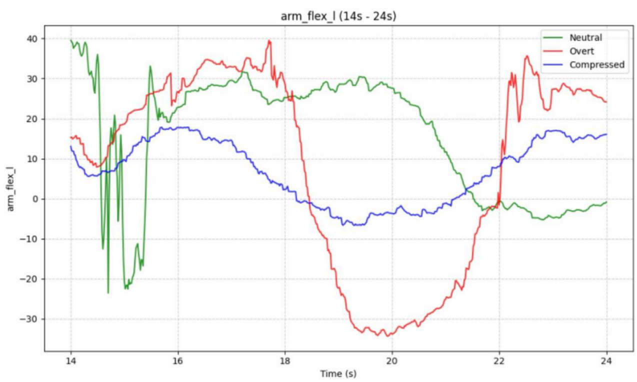

FIGURE 1. Left shoulder flexion-extension (arm_flex_l) in the 14-24 s phrase window across three intentional states

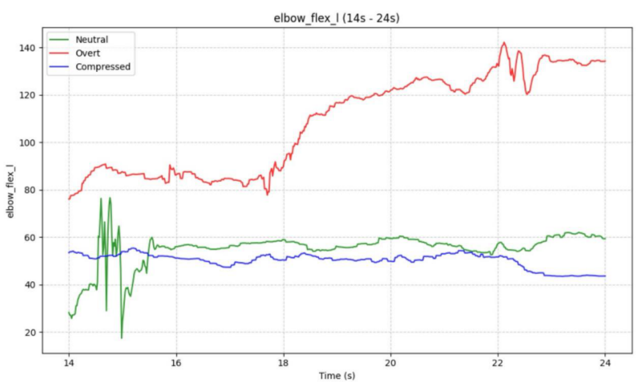

FIGURE 2. Left elbow flexion-extension (elbow_flex_l) in the 14-24 s phrase window across three intentional states

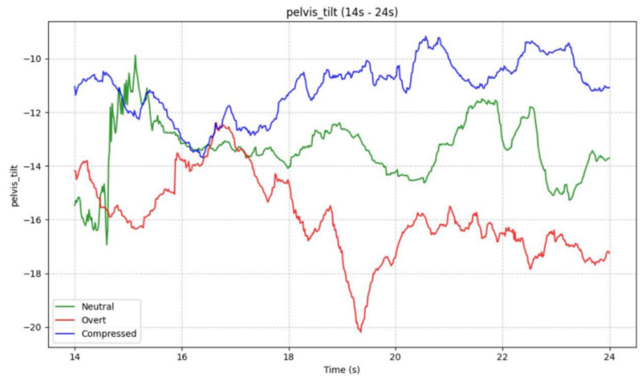

FIGURE 3. Pelvic tilt (pelvis_tilt) in the 14-24 s phrase window across three intentional states

C. ACOUSTIC DESCRIPTORS AND PHONATORY TRADE-OFFS

The acoustic descriptors tell a more ambivalent story. Table 3 shows that mean intensity varied only modestly across conditions: 52.39 dB in Neutral, 54.35 dB in Overt and 53.17 dB in Compressed. The central pitch corridor was also stable. The robust pitch span was 79.53 Hz in Neutral, 78.37 Hz in Overt and 79.59 Hz in Compressed. In this window, therefore, changing the gestural task did not greatly alter the central pitch range or average intensity.

TABLE 3. Acoustic descriptors and point-estimate acoustic contrasts relative to Neutral

Measure (14-24 s)	Neutral	Overt	Compressed	Overt - Neutral	Compressed - Neutral
Intensity mean (dB)	52.39	54.35	53.17	+1.96	+0.78
Intensity SD (dB)	11.00	12.29	10.86	+1.29	-0.14
Intensity range (dB)	42.83	44.34	39.37	+1.51	-3.46
Pitch mean (Hz)	173.17	168.24	174.77	-4.93	+1.60
Pitch SD (Hz)	34.69	26.39	28.85	-8.30	-5.84
Pitch robust span q95-q05 (Hz)	79.53	78.37	79.59	-1.16	+0.06
HNR mean (dB)	9.34	6.38	4.95	-2.96	-4.39
HNR SD (dB)	6.03	5.73	5.17	-0.30	-0.86

HNR showed the clearest acoustic difference. Mean HNR fell from 9.34 dB in Neutral to 6.38 dB in Overt and 4.95 dB in Compressed. HNR should not be treated as a direct grade of artistry or vocal quality. It is sensitive to phonation, noise, segmental material and measurement settings. Still, because the same phrase window and lyric content were compared, the drop in Compressed is a meaningful risk marker. It may indicate breathier phonation, additional constriction, or a more complex balance of acoustic energy and noise under the constraint of holding back movement.

Two explanations remain plausible. The first is loss of mechanical impulse. If high-impulse arm movement can physically contribute to respiratory-vocal dynamics, then reducing that movement may remove part of the support experienced in Overt. This explanation is consistent with gesture-speech physics, though those studies are based mainly on speech and laboratory tasks rather than legato Lied performance [10,11]. The second explanation is inhibition load. Compressed did not merely ask the singer to move less; it introduced a psychophysiological inhibition load—a dual-task constraint where the singer maintains the internal gestural impulse while actively suppressing its external manifestation. While the reflexive log describes this as a "felt need to stop the hands," this subjective report serves to contextualize the objective kinematic finding of altered pelvic tilt, suggesting that the effort of suppression may manifest as measurable axial bracing. It is also possible that both pathways operated together. The overt gesture may have supplied a clear spatial and kinetic pathway for the phrase, while the compressed version removed that pathway and introduced a secondary monitoring task. This would explain why intensity and central pitch span stayed broadly stable while HNR shifted. The singer may have preserved the musical contour and dynamic level, but with a less favourable noise-to-harmonic balance. In performance terms, this is exactly the kind of cost that can be missed by listening only for pitch and loudness.

D. REFLEXIVE TRIANGULATION: CUEING VERSUS INHIBITION

The reflexive log supports the inhibition-load explanation, but it does not prove it. In Neutral, the singer noted: "No deliberate gesture. Aim: keep body released. Focus on coordination, not acting." In Overt, the report was permissive: "The arm path helped the sound go forward, then gather back with the phrase energy." In compressed, the language changed: "Chest pressure/impulse; shoulders slightly stiff from stopping the hands. Hard to keep it loose." These entries match the acoustic pattern closely enough to guide interpretation. Overt felt like an external cue; Compressed felt like a contained impulse with a cost.

This is where Laban-informed thinking is useful, provided it is not overextended. Effort vocabulary helps name the difference between an outward gesture and a contained impulse, but it cannot by itself predict acoustic outcome. In this case, Compressed may have preserved some informational value as an internal cue while losing the physical impulse of Overt and adding a layer of inhibition. The term micro-movement therefore needs careful use. A micro-movement is not automatically refined, efficient or expressive. In this protocol, a functional micro-movement is kinematically distinguished from simple suppression by a reduction in distal limb Range of Motion (ROM) while maintaining or redistributing the phrase's impulse into axial postural adjustments. Without this redistribution, the movement may be categorized as 'over-compressed,' where the impulse is simply blocked rather than distilled. To ensure logical consistency, a distinction must be made between 'healthy economy'—where kinetic energy is efficiently internalized—and 'maladaptive bracing.' The current results suggest that without expert intervention, the 'Compressed' state risks becoming an 'Inhibited' state, where the effort of suppression manifests as the observed shoulder stiffness and acoustic degradation. The log also helps avoid a common error in interpreting movement data. A smaller trace is not always a calmer body. It can mean that the performer has solved the task economically, but it can also mean that motion has been blocked. Without the reflexive log, the Compressed condition might be described simply as successful visual reduction. With the log, it becomes a more complex case: visible economy was achieved, but the singer experienced a stopping action in the shoulders and chest.

E. METHODOLOGICAL IMPLICATIONS AND LIMITS

The findings clarify the visual-acoustic trade-off. From the audience side, Compressed is closer to art-song economy because it lowers visible arm action. From the singer's side, however, that economy may be bought at the cost of additional holding. The result does not mean that overt movement is preferable in performance. It means that a rehearsal gesture should not be compressed by simply suppressing its visible part. The performer needs to test whether the smaller version still supports phrase organisation or whether it has become a task of self-interruption.

The methodological contribution is the diagnostic protocol rather than the particular numerical outcome. While the findings from this single 10-second window are idiographic and cannot be generalized to all performers, they serve as a proof-of-concept for the measurement framework. The scientific weight of this study lies in its repeatable structure, which allows other performer-researchers to build a broader evidence base through collective application of the protocol. A singer can identify a phrase window, record Neutral/Overt/Compressed variants, extract a small number of kinematic and acoustic measures, and add a short reflexive log. The value lies in triangulation. Kinematic data show where movement is redistributed. Acoustic data show whether selected sound descriptors remain stable or shift. Reflexive notes show whether a take was experienced as released, cued, inhibited or over-managed. No stream is sufficient alone.

This protocol is intentionally small. It does not require a full motion-capture laboratory, a large corpus or complex modelling before it becomes useful in practice. For a performer-researcher, that matters. The method can be repeated across different phrase windows, repertoire types and pedagogical questions. It can also help teachers and singers move beyond general advice such as use your body or do less. The more useful question is local: what does this phrase ask the body to do, and what happens when that action is enlarged, contained or suppressed?

Several limitations must be kept explicit. The study involves one singer and one phrase window, so it cannot establish general gesture efficacy. The order was fixed, not counterbalanced. OpenCap is accessible and useful for gross kinematic organisation, but it cannot resolve all fine vocal-body adjustments. Praat descriptors are informative but do not replace expert listening, score analysis or physiological measurement. HNR in particular should be interpreted as a descriptive signal, not as a direct diagnosis. The reliance on HNR is a known limitation, as it is sensitive to recording environments and spectral fluctuations. To achieve a more robust analysis of vocal cost and breathiness, future studies implementing this protocol should incorporate Cepstral Peak Prominence (CPP), which offers superior reliability across different phonetic contexts compared to traditional time-domain measures. Finally, the study does not include audience data, so claims about visual reception remain grounded in the literature rather than measured in this pilot.

Even with these limits, the pilot answers a concrete performer problem. It shows that the move from rehearsal gesture to performance micro-movement is not a simple scaling operation. It can reorganise the body and alter selected acoustic descriptors. For art song, where the stage picture often demands economy, this is a useful warning. The task is not to eliminate gesture, nor to import large rehearsal movement onto the stage. The task is to find a smaller action that still carries direction without turning restraint into bracing.

IV. CONCLUSION

This article condensed a gesture-centred practice-led project into a focused pilot study of rehearsal-to-stage translation in art song. A single singer-researcher performed the same phrase window from Wolf's Ganymed under Neutral, Overt and Compressed conditions, combining OpenCap kinematics, Praat acoustics and reflexive logging. The central finding is that Compressed did not behave as a simple reduced copy of Overt. It reduced visible upper-limb range, especially at the elbow, but it also altered pelvic/postural organisation and coincided with the lowest mean HNR of the three conditions.

The most defensible interpretation is a trade-off. Overt gesture can work as an exploratory rehearsal cue and may carry mechanical as well as attentional benefits. Compressed gesture can make that impulse more compatible with art-song economy, but it may also remove useful physical impulse or add inhibition load. The study does not prove which mechanism dominates. It does show where the risk appears and how a singerresearcher can document it.

For performers and teachers, the practical implication is clear: do not assume that a covert version of a rehearsal gesture preserves the same function. Test it in a phrase window. Ask where the movement relocates, whether the sound remains stable, and whether the singer feels cued or braced. Future work should repeat the protocol across additional singers, songs, phrase types and counterbalanced condition orders, and should combine acoustic descriptors with expert listening or audience-perception data. The contribution of the present study is therefore modest but usable: a disciplined way to examine how visible economy, internal cueing and vocal cost interact in art song performance.

AUTHOR CONTRIBUTIONS

Yixiao Cao designed, collected and analyzed the data, and drafted the manuscript. Yixiao Cao conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Yixiao Cao participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] M. Leman, Embodied Music Cognition and Mediation Technology. Cambridge, MA, USA: MIT Press, 2008, DOI: 10.7551/mitpress/7476.001.0001.
[2] A. Cox, "Music and Embodied Cognition: Listening, Moving, Feeling, and Thinking," Bloomington, IN: Indiana University Press; 2016, DOI: 10.2307/j.ctt200610s.
[3] L. Peultier-Celli, M. Audouin, C. Beyaert, and P. Perrin, "Postural control in lyric singers," Journal of Voice, vol. 36, no. 1, pp. 141.e11-141.e17, 2022, DOI: 10.1016/j.jvoice.2020.04.019.
[4] S. D'Amario, S. Ternstrom, W. Goebl, and L. Bishop, "Body motion of choral singers," Frontiers in Psychology, vol. 14, Art. no. 1220904, 2023, DOI: 10.3389/fpsyg.2023.1220904.
[5] A. Jesse and D. W. Massaro, "Seeing a singer helps comprehension of the song's lyrics," Psychonomic Bulletin & Review, vol. 17, no. 3, pp. 323-328, 2010, DOI: 10.3758/PBR.17.3.323.
[6] C. J. Tsay, "Sight over sound in the judgment of music performance," Proceedings of the National Academy of Sciences of the United States of America, vol. 110, no. 36, pp. 14580-14585, 2013, DOI: 10.1073/pnas.1221454110.
[7] N. Moura, P. Fonseca, J. P. Vilas-Boas, and S. Serra, "Increased body movement equals better performance? Not always! Musical style determines motion degree perceived as optimal in music performance," Psychological Research, vol. 88, pp. 1314-1330, 2024, DOI: 10.1007/s00426-024-01928-x.
[8] J. W. Davidson, "The role of the body in the production and perception of solo vocal performance: A case study of Annie Lennox," Musicae Scientiae, vol. 5, no. 2, pp. 235-256, 2001, DOI: 10.1177/102986490100500206.
[9] M. C. Brunkan and J. Bowers, "Singing with gesture: Acoustic and perceptual measures of solo singers," Journal of Voice, vol. 35, no. 2, pp. 325.e17-325.e22, 2021, DOI: 10.1016/j.jvoice.2019.08.029.
[10] W. Pouw, S. J. Harrison, and J. A. Dixon, "Gesture-speech physics: The biomechanical basis for the emergence of gesture-speech synchrony," Journal of Experimental Psychology: General, vol. 149, no. 2, pp. 391-404, 2020, DOI: 10.1037/xge0000646.
[11] W. Pouw, S. J. Harrison, N. Esteve-Gibert, and J. A. Dixon, "Energy flows in gesture-speech physics: The respiratory-vocal system and its coupling with hand gestures," The Journal of the Acoustical Society of America, vol. 148, no. 3, pp. 1231-1247, 2020, DOI: 10.1121/10.0001730.
[12] L. Pearson and W. Pouw, "Gesture-vocal coupling in Karnatak music performance: A neuro-bodily distributed aesthetic entanglement," Annals of the New York Academy of Sciences, vol. 1515, no. 1, pp. 219-236, 2022, DOI: 10.1111/nyas.14806.
[13] R. Laban and F. C. Lawrence, "Effort: Economy in Body Movement," 2nd ed. London, UK: Macdonald & Evans; 1974.
[14] V. Ewan and K. Sagovsky, "Laban's Efforts in Action: A Movement Handbook for Actors with Online Video Resources," London, UK: Methuen Drama; 2019, DOI: 10.5040/9781472527325.
[15] R. P. Tsachor and T. Shafir, "A somatic movement approach to fostering emotional resiliency through Laban Movement Analysis," Frontiers in Human Neuroscience, vol. 11, pp. 410, 2017, DOI: 10.3389/fnhum.2017.00410.

[16] C. Wu, D. Davaasuren, T. Shafir, R. Tsachor, and J. Z. Wang, "Bodily expressed emotion understanding through integrating Laban movement analysis," Patterns, vol. 4, no. 10, Art. no. 100816, 2023, DOI: 10.1016/j.patter.2023.100816.
[17] R. I. Godoy, M. Leman, and editors, "Musical Gestures: Sound, Movement, and Meaning," New York, NY: Routledge; 2010, DOI: 10.4324/9780203863411.
[18] R. I. Godoy, "Coarticulated gestural-sonic objects in music," in Gritten A, King E, editors. New Perspectives on Music and Gesture. Farnham, UK: Ashgate, 2011, pp. 67-82.
[19] S. Dahl and A. Friberg, "Visual perception of expressiveness in musicians' body movements," Music Perception, vol. 24, no. 5, pp. 433-454, 2007, DOI: 10.1525/mp.2007.24.5.433.
[20] J. W. Davidson, "Bodily movement and facial actions in expressive musical performance by solo and duo instrumentalists: Two distinctive case studies," Psychology of Music, vol. 40, no. 5, pp. 595-623, 2012, DOI: 10.1177/0305735612449896.
[21] M. A. Smart, "Mimomania: Music and Gesture in Nineteenth-Century Opera," Berkeley, CA: University of California Press; 2004, DOI: 10.1525/california/9780520239951.001.0001.
[22] S. D. Uhlrich, A. Falisse, L. Kidzinski, J. Muccini, M. Ko, A. S. Chaudhari, S. L. Delp, and J. L. Hicks, "OpenCap: Human movement dynamics from smartphone videos," PLOS Computational Biology, vol. 19, no. 10, Art. no. e1011462, 2023, DOI: 10.1371/journal.pcbi.1011462.
[23] P. Boersma and D. Weenink, "Praat: Doing phonetics by computer [Computer software]," Amsterdam, The Netherlands: University of Amsterdam; 2024, [Online]. Available: https://www.fon.hum.uva.nl/praat/.