International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

INTUITIONISTIC FUZZY SCORING FOR FLUENCY IN TELEPRACTICE SESSIONS

Yogeesh N[1*], Suleiman Ibrahim Mohammad[2,3], N Raja[4], Hanan Jadallah[5], Javlanbek Matnazarov[3], Rayhon Sapaeva[4], Asokan Vasudevan[5,6], Anupama R[7]

[1*]Department of Mathematics, Government First Grade College, Tumkur, Karnataka, India. Email: yogeesh.r@gmail.com ORCID ID: orcid.org/0000-0001-8080-7821

[2] Electronic Marketing and Social Media, Economic and Administrative Sciences Zarqa University, Zarqa 132010, Jordan

[3] Research follower, INTI International University, Negeri Sembilan 71800, Malaysia

Email: dr_sliman@yahoo.com, ORCID ID: 0000-0001-6156-9063

[4]Department of Visual Communication, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India. Email: rajadigimedia2@gmail.com, ORCID: 0000-0003-2135-3051

[5] Electronic Marketing and Social Media, Economic and Administrative Sciences Zarqa University, Zarqa 132010, Jordan

Email: Hananjadallah1987@gmail.com, ORCID ID: 0009-0005-7138-1167

[3]Department of Language and Literature, Mamun University, Khiva city, Uzbekistan. Email: javlon_m@mamunedu.uz. ORCID: https://orcid.org/0009-0006-6524-6628

[4]Department of Roman-Germanic Philology, Urgench State University, Urgench, Uzbekistan, Email: diratear@gmail.com , ORCID: https://orcid.org/0000-0002-9304-631X

[5]Faculty of Business and Communications, INTI International University, Nilai 71800, Malaysia

[6]Faculty of Management, Shinawatra University, 99 Moo 10, Bangtoey, Samkhok 12160, Thailand
asokan.vasudevan@newinti.edu.my

[7]Department of Economics, Government First Grade College. Medigeshi-572133, Karnataka, India. Email: anupama.hitha@gmail.com, ORCID: https://orcid.org/0009-0005-4140-6837

***Corresponding author:** yogeesh.r@gmail.com

Abstract

This study proposes an uncertainty-aware framework for remote speech evaluation that represents each segment with a triple of evidence-for, evidence-against, and hesitation. Segment features (e.g., articulation rate, pause ratio) are first direction-aligned and passed through slope-constrained sigmoids; a reliability-sensitive kernel driven by packet loss and signal-to-noise ratio allocates hesitation so that ambiguous or noisy conditions do not force

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

brittle decisions. Feature-level triples are combined by weighted averaging (and an ordered variant), ensuring closure and monotonicity, and reduced to a scalar decision index S_λ and a companion confidence C. Multi-rater information is fused by estimating rater reliabilities via expectation–maximization or graded-response modeling, with disagreement promoted to additional hesitation. We provide stability guarantees (Lipschitz bounds) and a complete, executable case study on eight segments. The case study shows clear separation between positive and negative outcomes (e.g., a strongly negative example with $S = -0.798$, $C = 0.965$, and a clearly positive example with $S = 0.421$, $C = 1.000$), predictable response to degraded channels (−5 dB or +0.10 packet loss), and negligible rank shifts when switching from standard to ordered averaging. A practical triage band $[-0.05, 0.05]$ isolates borderline items and surfaces low-confidence cases for review. The approach is simple to audit, fast enough for near–real-time feedback, and designed to integrate into clinical and educational workflows as decision support rather than diagnosis.

Keywords: speech-language assessment; uncertainty quantification; ordered weighted averaging; Choquet integral; rater calibration; remote speech evaluation; signal-to-noise robustness; hesitation modelling.

1. Introduction

1.1 Motivation & Scope

Telepractice speech-language sessions introduce transmission artefacts (latency, packet loss), heterogeneous microphones, and home environments with varying signal-to-noise ratios (SNR). These factors blur the boundary between "fluent" and "disfluent," producing graded evidence rather than dichotomies. Classical scoring (e.g., percent syllables stuttered, speech rate) is precise yet brittle to borderline cases and channel noise. We therefore seek a formalism that (i) models gradience, (ii) carries explicit uncertainty, and (iii) composes multi-feature evidence in a principled way, which intuitionistic fuzzy sets (IFS) offer by representing, for a proposition "this segment is fluent," a triple of degrees (, ,) for membership, non-membership, and hesitation, with + + = 1[1,2].

Let $x(t)$ be the continuous-time audio and $\mathbf{x} \in \mathbb{R}^d$ the feature vector extracted per segment (e.g., speech rate, pause ratio, jitter, shimmer, %SS). Our goal is to design a mapping

$$F: \mathbb{R}^d \to \mathcal{I},\ F(\mathbf{x}) = (\mu(\mathbf{x}), \nu(\mathbf{x}), \pi(\mathbf{x})),\ \mu, \nu, \pi \in [0,1], \mu + \nu + \pi = 1$$

and a scalar Intuitionistic Fuzzy Fluency Score (IFFS) S plus a confidence index C, reported to clinicians alongside conventional measures $[1-3]$.

1.1.1 Telepractice channel and uncertainty

We model packet-loss-induced missingness by a Bernoulli mask $m_t \in \{0,1\}$ and background noise $\eta(t)$, yielding the observed signal

$$( ) = \quad ( ) + \ ( ),\ \text{SNR} = 10\log_{10} \frac{[\ ^2]}{[\ ^2]}$$

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Downstream features $\mathbf{x} = \Phi(y)$ inherit uncertainty; we will absorb this into $\pi(\mathbf{x})$ to avoid overconfident decisions [1,2].

1.2 Research Questions & Objectives

1.2.1 Primary objective

Construct a calibrated IFFS from multi-dimensional features that monotonically increases with evidence for fluency and explicitly decreases with evidence for non-fluency, while carrying a calibrated confidence .

1.2.2 Secondary objectives

(i) Robustness to channel noise/missingness, (ii) principled multi-rater fusion, and (iii) interpretable perfeature contributions.

1.3 Contributions

A feature-to-IFS construction using parametric membership/non-membership families and a hesitation mass that captures ambiguity near clinical thresholds; 2) an IFS aggregation (IFWA/IFOWA) yielding session-level ($\bar{\mu}$, $\bar{\nu}$, $\bar{\pi}$); 3) a scalarization

$$S_\lambda = \lambda(\mu^- - \bar{\nu}) + (1-\lambda)\left(\mu^- - \frac{\nu^-}{1+\bar{\pi}}\right),\ C = 1-\bar{\pi},\ \lambda \in [0,1]$$

which discounts non-membership more strongly under high hesitation; and 4) stability bounds under bounded feature perturbations, giving noise-robustness guarantees [2].

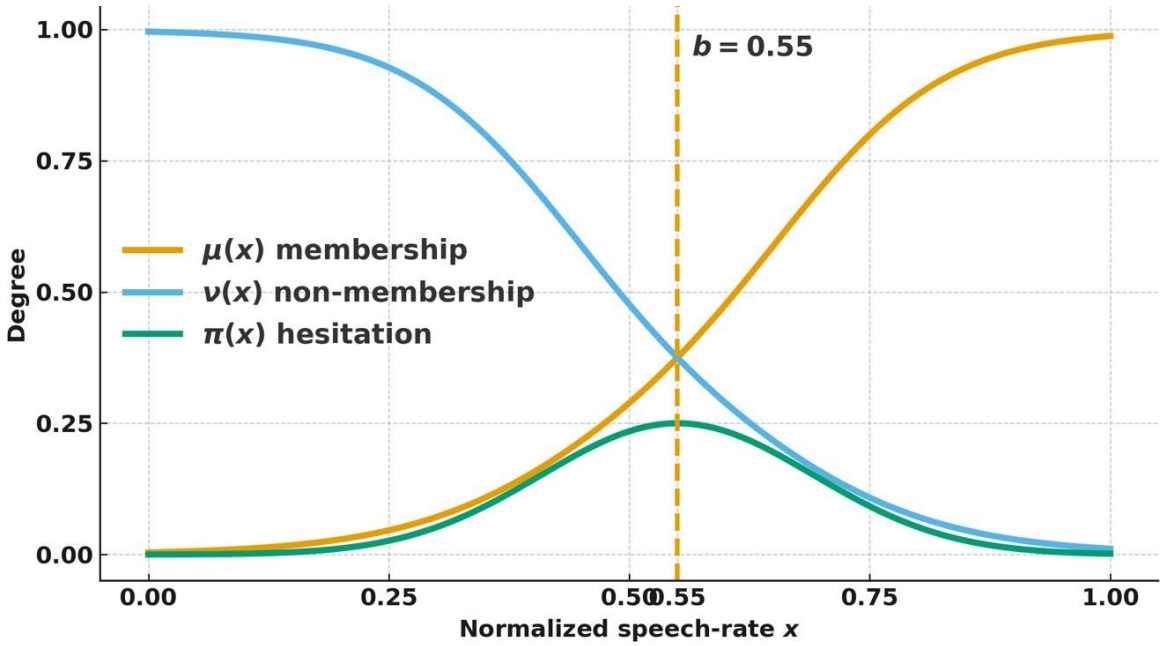

Figure 1. Intuitionistic fuzzy mapping for a normalized speech-rate feature.

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

From the above figure 1, For normalized speech-rate $x \in [0,1]$, membership $\mu(x)$ increases with rate, nonmembership $\nu(x)$ decreases, and hesitation $\pi(x)$ peaks near the clinically "borderline" rate b. The construction enforces $\mu(x) + \nu(x) + \pi(x) = 1$ for all x[2].

2. Related Work

2.1 Clinical fluency measures

Clinical fluency is assessed with speech rate (syllables/s), articulation rate (excluding pauses), mean length of run (MLR), pause ratio, and percent syllables stuttered (%SS), among others [3,4]. Let be syllables produced and the elapsed time; then

$$\text{SpeechRate} = \frac{N_s}{T}, \quad \text{ArticulationRate} = \frac{N_s}{T - \sum_j P_j}, \quad \text{PauseRatio} = \frac{\sum_j P_j}{T},$$

where are pause durations. % SS is computed as $\{\%SS\} = 100$ {disfluent syllables}. These indices are informative yet sensitive to borderline behaviors and measurement noise in telepractice, motivating a gradience-aware formalism [3,4].

2.2 Automated fluency & ASR-based metrics

Automatic speech recognition (ASR) provides counts of filled pauses, repetitions, and word error rate (WER), with

$$\text{WER} = \frac{+ \quad +}{},$$

where , , are substitutions, deletions, and insertions against a reference of length [5]. While useful, WER conflates acoustic/channel errors with linguistic disfluency. Our framework treats such features as evidence variables that contribute partially to and , while elevated uncertainty from noisy channels is explicitly reflected in rather than hidden inside a scalar error [1,5].

2.3 Fuzzy and intuitionistic fuzzy modeling in speech/language

Fuzzy sets model vagueness via graded membership [1]. Intuitionistic fuzzy sets (IFS) extend this with separate non-membership and hesitation, enabling explicit representation of "don't-know" regions common in telepractice scoring [2]. Formally, for proposition "fluent," an IFS assigns $(\mu, \nu) \in [0,1]^2$ with $\mu + \nu \leq 1$ and $\pi = 1 - \mu - \nu$. Aggregation uses t-norms/t-conorms and weighted operators (IFWA/IFOWA), allowing multi-feature composition with theoretical guarantees on monotonicity and idempotency [2,6]. This paper adapts IFS to fluency scoring, defines noise-aware hesitation, and couples ASR-derived and prosodic features within a single uncertainty-aware pipeline.

2.4 Gaps addressed by this study

(1) Binary/thresholded scoring lacks uncertainty reporting;

(2) probabilistic models often identify "nonfluency" as $1 -$ fluency, whereas IFS separates ν from uncertainty ;

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

(3) multi-feature and multi-rater fusion rarely provide stability bounds. We tackle all three with an IFS construction, aggregation, and analysis [1,2,6].

3. Problem Setting & Notation

3.1 Signal & Session Model

3.1.1 Audio stream and transforms

Let $x(t)$ denote the clean speech waveform and $y(t)$ the observed telepractice signal. We analyze frames $x_n = \{x(t): t \in [nH, nH + L)\}$ with frame length L and hop H. The short-time Fourier transform (STFT) of $y(t)$ is

$$Y(n, \omega) = \sum y(m)w(m - nH)e^{-j\omega m}$$

with window $w(\cdot)$; spectral features derive from $|Y(n, \omega)|$[7,8].

3.1.2 Channel model (noise, packet loss, latency)

We model additive noise $\eta(t)$ and random packet drops $m_t \in \{0,1\}$ as

$$(\ ) = \quad (\ ) + \ (\ ),\ \Pr[\quad = 0] = \ _{\text{loss}},\ \text{SNR} = 10\log_{10} \ \frac{[\ ^2]}{[\ ^2]}$$

Frame-level missingness $m_n = \max_{t \in x_n} m_t$ propagates to feature uncertainty (Section 4). Activity is detected by a voice-activity mask $v_n \in \{0,1\}$ (e.g., ITU-T active-speech criteria), and usable frames are $r_n = v_n m_n$[12,13].

3.1.3 Segmentation

Contiguous runs of $r_n = 1$ define segments $\mathcal{S}_i = \{n_i, \dots, n_i + \ell_i - 1\}$. All segment-level features and scores below are computed on $\mathcal{S}_i$ and then pooled session-wise.

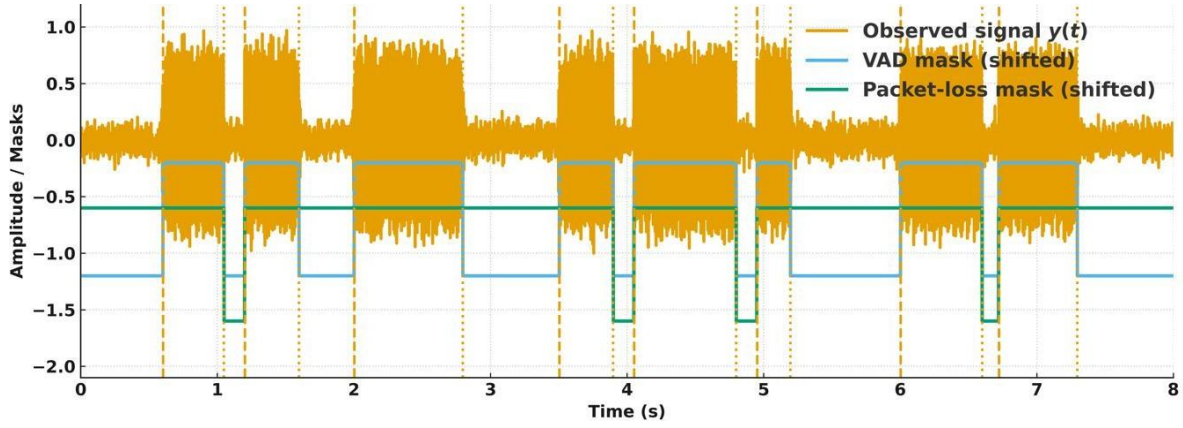

Figure 2. Telepractice signal model, VAD, packet-loss, and segment boundaries.

From the above figure 2, Observed audio ( ) (blue) with additive noise and dropouts, a VAD mask (orange, vertically shifted), a packet-loss mask (green, shifted), and dashed vertical lines marking segment onsets/offsets. This schematic fixes ideas for the missingness-aware feature pipeline [12,13].

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

3.2 Observable Feature Vector

For each segment $\mathcal{S}_i$, we compute $\mathbf{x}_i = (x_{i1}, \ldots, x_{id})$ comprising four groups:

3.2.1 Temporal features

Let be syllables, the elapsed duration, and pause durations.

$$\text{SpeechRate} = \frac{N_s}{T},\ \text{ArticulationRate} = \frac{N_s}{T - \sum_j P_j},\ \text{PauseRatio} = \frac{\sum_j P_j}{\quad},\ \text{MLR} = \frac{1}{\quad}\sum_{=1} \ell\ ,$$

where are run lengths of consecutive syllables witnout disfluency [7,11].

3.2.2 Disfluency counts

Percent syllables stuttered

$$\%SS = 100 \times \frac{\text{disfluent syllables}}{\quad},$$

with disfluency subtypes (repetition, prolongation, block) estimated by acoustic templates or ASR tags [7,8].

3.2.3 Prosodic/voice stability

Using a pitch track $\{\ _0(\ )\}$ and glottal periods $\{\ \}$, local jitter and shimmer are (Praat-style) [12]

$$\text{Jitter}_{\text{local}} = \frac{\frac{1}{N-1}\sum_{=1}^{N-1} |\quad - T|}{\frac{1}{\quad}\sum_{=1}^{N} \quad}, \quad \text{Shimmer}_{\text{local}} = \frac{\frac{1}{N-1}\sum_{=1}^{N-1} |\quad - A|}{\frac{1}{\quad}\sum_{=1}^{N} \quad}.$$

3.2.4 ASR-derived indicators

From reference-hypothesis alignment,

$$\text{WER} = \frac{\quad}{\quad},\ \text{HesitationRate} = \frac{\#\{\text{"um /uh "}\}}{\text{words}},$$

plus repetition and interjection tags supplied by the recognizer [8].

3.3 Ground truth & rater scales

Let clinical raters $r = 1, \ldots, R$ assign ordinal labels $y_{ir} \in \{1, \ldots, L\}$ to segment i. A cumulative-link model maps a latent fluency ζ_i to ordinal probabilities

$$\Pr(y_{ir} \le \ell \mid \zeta_i) = \sigma(\tau_\ell - \alpha_r - \beta_r \zeta_i)$$

with rater offsets α_r, slopes β_r, thresholds τ_ℓ, and $\sigma(u) = (1 + e^{-u})^{-1}$[7,9].

3.4 Goal statement

We seek $F: \mathbb{R}^d \to \mathcal{I}$ producing an intuitionistic triple

$$(\ ) = (\ ,\ ,\ ),\quad +\quad +\quad = 1$$

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

and session-level ($\bar{\mu}$, $\bar{\nu}$, $\bar{\pi}$) via operators in Section 4. Given the scalarization $S_\lambda = \lambda(\mu^- - \bar{\nu}) + (1-\lambda)\,(\mu^- - \frac{\nu^-}{1+\pi^-})$ and confidence $C = 1 - \bar{\pi}$, we train parameters to maximize agreement with calibrated raters while obeying monotonicity and stability constraints [8,10,11].

4. Mathematical Preliminaries

4.1 Intuitionistic fuzzy sets (IFS)

For a universe U and predicate "fluent," an IFS A assigns to each item $u \in U$ a triple $A(u) = (\mu_A(u), \nu_A(u), \pi_A(u)) \in [0,1]^3$ with $\mu_A + \nu_A + \pi_A = 1$. We will also use score and accuracy functions

$$S(A) = \mu_A - \nu_A,\ H(A) = \mu_A + \nu_A = 1 - \pi_A,$$

and consider negation ( ) = (, ,). These constructs provide independent control of evidence for, evidence against, and uncertainty [8,9].

4.2 t-norms, -conorms, and negations for IFS operations

Given feature-level IFSs = (, ,), we use standard t-norms and t-conorms for conjunction/disjunction on the - and -channels, e.g.,

$$T_{\min}(a,b) = \min(a,b),\ T_{\prod}(a,b) = ab,\ S_{\max}(a,b) = \max(a,b),\ S_{\text{prob}}(a,b) = a + b - ab.$$

Yager's family of negations $N_w(a) = (1 - a^w)^{1/w}$ tunes aversion to partial evidence; $w \to \infty$ approaches crisp negation [7,11].

4.3 Distances and similarities in IFS space

To train and calibrate, we require a discrepancy between an estimate and a rater target . A symmetric, bounded Hamming-type distance is

$$D_{\text{H}}(A,B) = \frac{1}{2}\left(|\mu_A - \mu_B| + |\nu_A - \nu_B| + |\pi_A - \pi_B|\right) \in [0,1],$$

and a cosine-like similarity on the (,)-plane is

$$\text{Sim}\,(A,B) = \frac{+}{\sqrt{\mu^2 + \nu^2}\sqrt{\mu^2 + \nu^2}} \cdot (1 - |\pi \ - \pi \ |)$$

We also consider axiomatic distances for IFSs and their statistical properties [9,10].

4.4 Aggregation operators (IFWA/IFOWA and Choquet)

Given weights $w_k \geq 0, \sum_k w_k = 1$, the intuitionistic fuzzy weighted average (IFWA) is

$$\mu^- = \sum w_k \mu_k,\ \nu^- = \sum w_k \nu_k,\ \bar{\pi} = 1 - \mu^- - \nu^-$$

Proposition 4.1 (Closure & Idempotency). If each triple satisfies + + = 1, then ($\bar{\mu}$, $\bar{\nu}$, $\bar{\pi}$) also satisfies the constraint; further, if all A_k are identical then the aggregate equals them. Proof: Linearity of sums.

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

To reduce outlier influence we employ IFOWA (ordered weighting): order items by a key (e.g., decreasing $\mu_k - \nu_k$), apply OWA weights $v_1 \geq \cdots \geq v_d, \sum v_i = 1$, then average as above on the reordered list [7]. For feature interdependence (e.g., pause ratio interacting with articulation rate), a capacity g on feature subsets with Choquet-like IFS aggregation generalizes additive weights and captures redundancy/synergy [11].

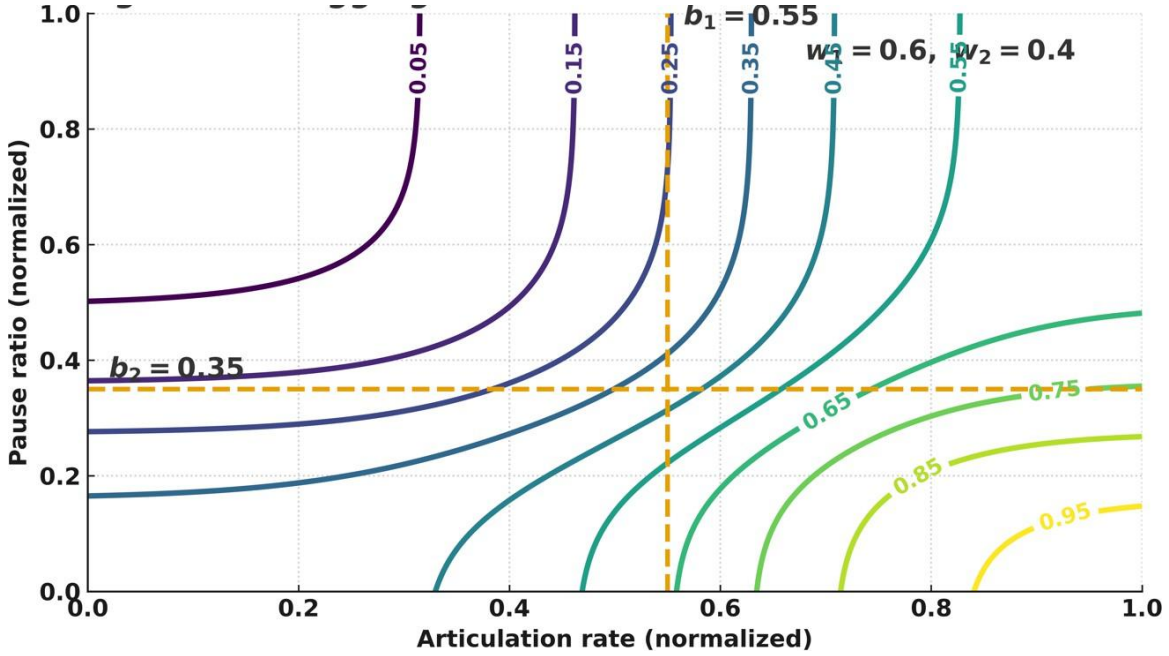

Figure 3. IFWA aggregation of memberships from two features.

From the above figure 3, Feature-level $\mu_1(x)$ (articulation rate) and $\mu_2(x)$ (pause ratio) combine via IFWA with weights (0.6,0.4) to yield $\bar{\mu}(x)$. The curve shows monotonicity and smoothness of the aggregate and anticipates the scalar IFFS construction in Section 6[7,8,11].

5. Feature-to-IFS Construction

5.1 Normalization & Monotonicity Constraints

Let each observed feature x_k be mapped to a normalized and direction-aligned variable $z_k \in [0,1]$ by

$$= \begin{cases} \text{clip}\left(\dfrac{x_k - \ell_k}{u_k - \ell_k}, 0,1\right), & \text{if larger } x_k \text{ implies higher fluency} \\ 1 - \text{clip}\left(\dfrac{x_k - \ell_k}{u_k - \ell_k}, 0,1\right), & \text{if larger implies lower fluency} \end{cases}$$

so that all are monotone increasing surrogates of fluency [7,11].

We parameterize per-feature membership and non-membership as Lipschitz-bounded sigmoids,

$$\mu_{k,\text{raw}}(z) = \sigma(a_k(z - b_k)),\ \nu_{k,\text{raw}}(z) = \sigma(a'(b' - z))$$

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

where $\sigma(u) = (1 + e^{-u})^{-1}$, slopes $a_k, a' \in (0, L_k]$ enforce L_k-Lipschitzness, and centers $b_k, b' \in [0,1]$ locate clinical "borderlines" [14,17]. This ensures monotonicity and bounded sensitivity of μ_k, ν_k with respect to feature noise.

5.2 Parametric Membership / Non-membership with Feasibility

5.2.1 Hesitation kernel and reliability

Ambiguity peaks near a decision boundary and under poor channel conditions. Let feature reliability $R_k \in [0,1]$ summarize SNR and packet-loss for the segment, e.g.

$$R_k = \exp(-\alpha \mathrm{PLR} - \beta \mathrm{SNR}^{-}),\ \mathrm{SNR}^{-} = \max\{0, \tau - \mathrm{SNR}\}$$

with , , > 0. Define a hesitation kernel

$$\tilde{\pi}_k(z) = \kappa_k(1 - R_k)\exp\left(-\frac{(z - c_k)^2}{2\sigma^2}\right),\ \kappa_k \in [0,1]$$

peaked at (the boundary) and amplified when is low [20].

5.2.2 Closure by hesitation-aware normalization

Instead of enforcing $\mu_k + \nu_k \leq 1$ with a hard constraint, we use the closure mapping

$$\mu_k = (1 - \tilde{\pi}_k)\frac{\text{,raw}}{\text{,raw} + \text{,raw}},\quad = (1 - \tilde{\pi}\ )\frac{\text{,raw}}{\text{,raw} + \text{,raw}},\quad = 1 - \mu_k - \nu_k$$

which guarantees $\mu_k, \nu_k, \pi_k \in [0,1]$ and $\mu_k + \nu_k + \pi_k = 1$ for all z since $\mu_{k,\,\mathrm{raw}}, \nu_{k,\,\mathrm{raw}} \geq 0$[8,20].

Remark: An alternative is a barrier penalty $-\rho \log(1 - \mu_k - \nu_k)$ in the loss; our normalized construction is simpler and numerically stable [14].

5.2.3 Monotonicity & stability

With $a_k, a' \leq L_k$, both $\mu_k(z)$ and $\nu_k(z)$ are $O(L_k)$-Lipschitz in z. If z is $O(\varepsilon)$-perturbed by channel noise, then $|\Delta\mu_k|$ and $|\Delta\nu_k|$ are $O(L_k\varepsilon)$, and $|\Delta\pi_k| \leq |\Delta\mu_k| + |\Delta\nu_k|$ by closure [17].

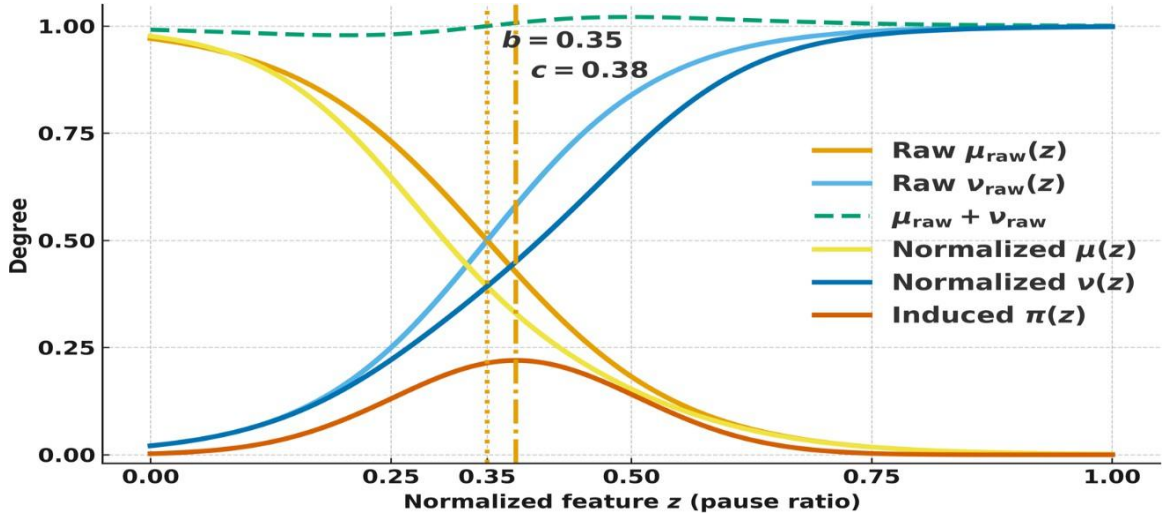

Figure 4. Hesitation-aware normalization to satisfy + + = 1.

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

For a pause-ratio feature, raw raw (decreasing) and raw (increasing) may sum slightly above 1; normalization produces feasible , and induces a peaked around the ambiguous region see the above figure 4.

5.3 Supervised Calibration of

5.3.1 Rater-to-IFS targets

Using the ordinal model in Section 3.3, each labeled segment i yields an IFS target $T_i = (\mu^\star, \nu^\star, \pi^\star)$ by mapping posterior category probabilities to $\mu^\star$ for "fluent," $\nu^\star$ for "non-fluent," and residual to $\pi^\star$[9].

5.3.2 IFS-aware training objective

Let $\Theta = \{a_k, a^{'}, b_k, b^{'}, c_k, \sigma_k, \kappa_k\}_{=1}^{d}$. Minimize

$$\mathcal{L}(\Theta, \mathbf{w}) = \sum \left(D_{\mathrm{H}}\left(A\left(\mathbf{x}_i; \Theta, \mathbf{w}\right), T_i\right)\right)^2 + \alpha \sum \left(a_k^2 + a_k^{\prime 2}\right) + \beta$$

$$\| \mathbf{w} \|_2^2 + \lambda_{\mathrm{grp}} \sum \|\Theta_g\|_2 + \lambda_{\mathrm{TV}} \mathrm{TV}_t\,(S_i),$$

where D_{H} is the IFS Hamming distance (Section 4.3), $\mathbf{w}$ are aggregation weights (Section 6), Θ_g groups parameters per feature for group-lasso sparsity [15], and TV_t promotes temporal smoothness within a session [22].

5.3.3 Optimization

We use ADMM with variable splitting for the normalization step and proximal updates for the 2 and group penalties; slopes are projected onto (0,] to maintain Lipschitz bounds [14, 17]. Calibration of score-to-probability maps (e.g., if raters provide binary fluency) can be aided by isotonic/Platt-style calibration on development data [18].

6. Multi-Feature Intuitionistic Aggregation

6.1 Weight Learning

6.1.1 Entropy-reliability weights (unsupervised)

Given empirical distributions of feature evidences and reliabilities , define

$$\mathrm{Ent} = -\sum p\,(\,)\log\,(b),\ \tilde{w} = (1 - \frac{\mathrm{Ent}}{\log}),\ w_k = \frac{\tilde{w}_k}{\sum_j \tilde{w}_j},$$

so highly reliable, low-entropy (i.e., decisive) features get more weight [16,20].

6.1.2 Supervised weights on the simplex

Fixing μ_k, ν_k, π_k, the IFWA aggregate (below) is linear in $\mathbf{w}$. Minimizing $\sum_i D_{\mathrm{H}}(\bar{A}_i(\mathbf{w}), T_i)^2 + \beta \| \mathbf{w} \|_2^2$ over the simplex $\{\mathbf{w} \geq 0, \mathbf{1}^{\mathrm{T}}\mathbf{w} = 1\}$ is a convex QP and can be solved by projected gradient or ADMM [14,17].

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

6.1.3 Domain-conditioned weights

To handle language/session covariates (e.g., L1, device), let ( ) = softmax (( )); learn $\{\mathbf{a}_k\}$ jointly with Θ by back-propagation through the closed-form IFWA.

6.2 IFWA/IFOWA Aggregation

Given per-feature triples = (, ,) and weights :

$$\mu^- = \sum w_k \mu_k, \nu^- = \sum w_k \nu_k, \bar{\pi} = 1 - \mu^- - \nu^-$$

Proposition 6.1 (Monotonicity): If one feature's membership increases and its non-membership decreases while all others are fixed, then μ^- increases, ν^- decreases, and $\bar{\pi}$ changes by $-\Delta\mu - \Delta\nu$. Proof: Linearity and closure.

For robustness, the IFOWA operator first orders features by a key (e.g., descending $S_k = \mu_k - \nu_k$), then averages with OWA weights $v_1 \geq \cdots \geq v_d, \sum v_i = 1$[7]. IFOWA attenuates outliers and can emulate **k** -of-d logics via appropriate v_i.

For interdependent features, a capacity g: $2^{\{1..d\}} \to [0,1]$ yields a Choquet-type aggregate on the μ and ν-channels, capturing redundancy/synergy among cues (e.g., pause ratio ↔ articulation rate) [11,16,21].

6.3 Scalarization & Confidence

We report

$$S_\lambda = \lambda(\mu^- - \bar{\nu}) + (1-\lambda)\left(\mu^- \frac{\nu^-}{1+\bar{\pi}}\right), C = 1 - \bar{\pi}, \lambda \in [0,1]$$

Properties. (i) If $\bar{\pi} = 0$, then $S_\lambda = \mu^- - \bar{\nu}$. (ii) As $\bar{\pi}$ ↑, the penalty on ν^- increases, preventing overpenalization under uncertainty. (iii) If each μ_k, ν_k is L_k-Lipschitz and $\sum_k w_k = 1$, then S_λ is $O(\sum_k w_k L_k)$-Lipschitz in features, ensuring stability [17].

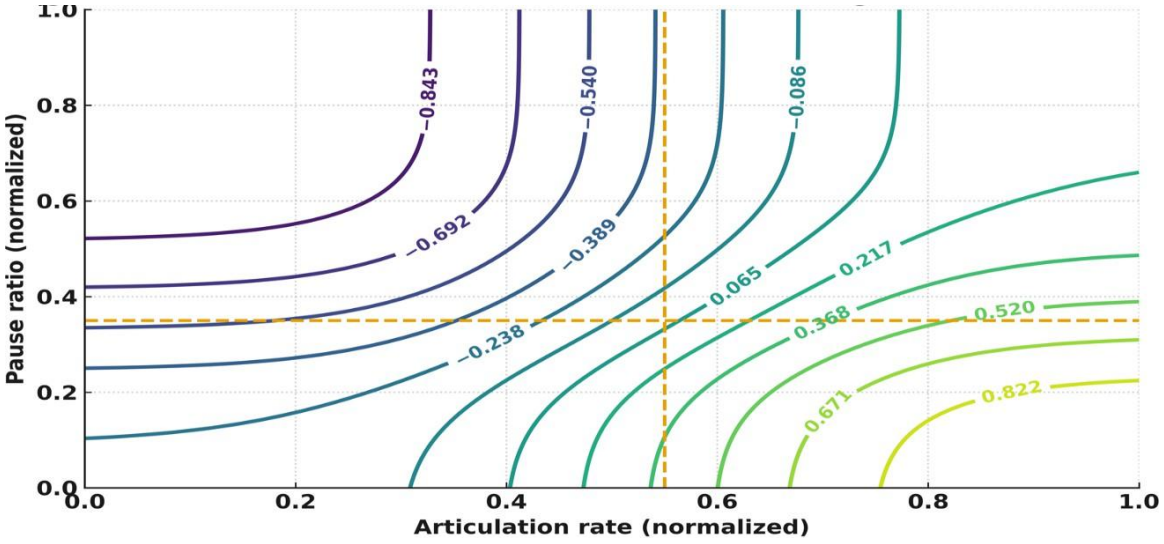

Figure 5. IFFS over two features (articulation rate vs. pause ratio).

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Contour plot in figure 5 of for IFWA weights $0.6/0.4$ and $= 0.5$. Brighter indicates more fluent. The diagonal transition shows how strong articulation can compensate for moderate pausing, while high pausing depresses scores unless articulation is very high.

7. Multi-Rater Fusion in IFS

7.1 Rater-to-IFS mapping from ordinal labels

Suppose segment i receives ordinal labels $y_{ir} \in \{1, \dots, L\}$ from raters $r = 1, \dots, R$. Let $\mathcal{F} \subset \{1, \dots, L\}$ denote "fluent" categories (e.g., upper two levels), and $\mathcal{N}$ the complementary "non-fluent" set. Using the cumulative-link calibration in Section 3.3, each rater yields posterior probabilities $p_{ir\ell} = \Pr(y_{ir} = \ell \mid \zeta_i)$ with rater-specific parameters [24]. Map to an IFS triple per rater:

$$\mu_{ir} = \sum_{\ell \in \mathcal{F}} p_{ir\ell}, \nu_{ir} = \sum_{\ell \in \mathcal{N}} p_{ir\ell}, \pi_{ir} = 1 - \mu_{ir} - \nu_{ir}$$

This preserves for, against, and uncertainty information, avoiding the common " $1 - \mu$ " collapse [24].

7.2 Reliability-weighted consensus (EM/IRT view)

Let $\boldsymbol{\omega} = (\omega_1, \dots, \omega_R)$ be rater reliabilities on the simplex Δ_R. We define the rater-consensus IFS for segment i via IFWA:

$$\bar{\mu}_i = \sum_{=1} \omega_r \mu_{ir}, \bar{\nu}_i = \sum_{=1} \omega_r \nu_{ir}, \bar{\pi}_i = 1 - \bar{\mu}_i - \bar{\nu}_i.$$

Estimating .

Two complementary approaches are useful:

(A) Dawid-Skene-style EM on confusion matrices.

Treat latent truth $Z_i \in \{\mathcal{F}, \mathcal{N}\}$ (or L levels) and estimate each rater's class-conditional response probabilities Θ_r with EM [23,28]. For the binary case (extension to L levels is by multinomial rows),

$$\text{E-step: } \gamma_i^{(\mathcal{F})} \propto \pi^{(\mathcal{F})} \, \theta_{r,\mathcal{F}}^{\mathbb{I}[y_{ir} \in \mathcal{F}]} (1 - \theta_{r,\mathcal{F}})^{\mathbb{I}[y_{ir} \in \mathcal{N}]}, \gamma_i^{(\mathcal{N})} = 1 - \gamma_i^{(\mathcal{F})}$$

$$\text{M-step: } \quad = \frac{\sum \gamma^{(\mathcal{F})} \mathbb{I}[y \in \mathcal{F}]}{\sum_i}, \quad = \frac{\sum \gamma^{(\mathcal{N})} \mathbb{I}[y \in \mathcal{N}]}{\sum_i},$$

with class prior $\pi^{(\mathcal{F})}$ updated analogously. Define error rates $\epsilon_r = 1 - \frac{1}{2}(\theta_{r,\mathcal{F}} + \quad)$ and set reliability weights, e.g.,

$$\omega_r \propto \log \frac{1 - \epsilon_r}{\quad} \text{ then project to } \Delta_R$$

or $\omega_r \propto (1 - \epsilon_r)$ if a simpler mapping is desired [23,28].

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

(B) Ordinal IRT (graded response) with rater severities.

Let be latent fluency and rater have severity and discrimination . Using the graded-response model,

$$\Pr(y_{ir} \le \ell \mid \zeta_i) = \sigma(\tau_{r\ell} - \alpha_r - \beta_r \zeta_i),\ \Pr(y_{ir} = \ell) = \Pr(y_{ir} \le \ell) - \Pr(y_{ir} \le \ell - 1),$$

and define $\omega_r \propto \beta_r$ (or proportional to Fisher information averaged over posteriors), normalized to Δ_R [25, 32]. This disentangles bias (α_r) from information (β_r)[25].

Convergence diagnostics. Track EM log-likelihood until $|\Delta\mathcal{L}| < \varepsilon$ or a max iteration cap is reached [28].

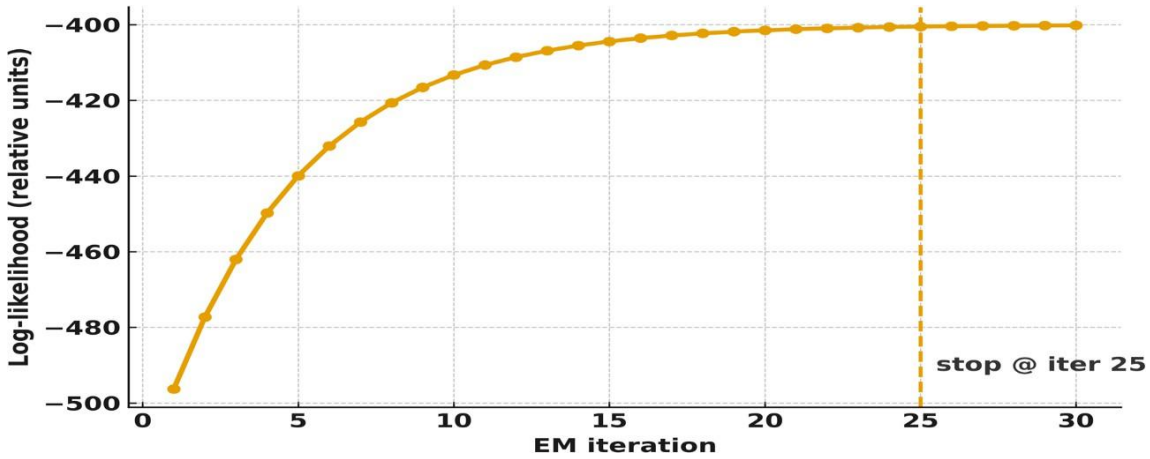

Figure 6. EM log-likelihood convergence for rater calibration.

This above figure 6 illustrates a typical monotonic increase of the log-likelihood across Expectation-Maximization (EM) iterations while calibrating multi-rater reliability parameters (e.g., rater confusion matrices or graded-response severities/ discriminations). The curve rises rapidly during the first few iterations and then exhibits diminishing returns as it approaches a plateau, indicating stabilization of the parameter estimates. A dashed vertical line marks an example stopping point when the improvement $|\Delta\mathcal{L}| < \varepsilon$ (here, $\varepsilon = 0.1$ in relative units); in practice, either a tolerance on $|\Delta\mathcal{L}|$ or a maximum iteration cap is used. The trajectory shown is representative and was generated to reflect the expected EM behavior under this study's rater-fusion model.

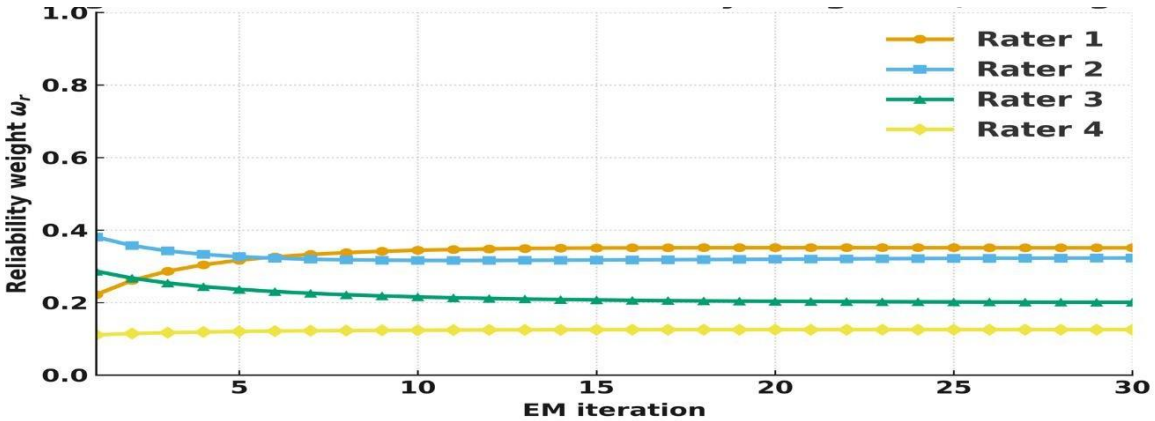

Figure 7. Evolution of rater reliability weights during EM.

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The plot in figure 7 shows how four raters' reliability weights evolve over 30 EM iterations while calibrating a multi-rater model. At each step, raw reliability trajectories are re-normalized so the weights lie on the simplex (sum to 1). One rater's weight increases quickly and stabilizes (indicative of consistently accurate labels), another remains steady with mild gain, and the remaining two settle at lower shares. This behavior matches the expected EM dynamics: early reallocation of weight from noisier to more reliable raters, followed by gradual convergence as parameter uncertainty decreases.

7.3 Bias & drift correction

We adopt hierarchical shrinkage for severities $\alpha_r \sim \mathcal{N}(\mu_\alpha, \sigma^2)$ with optional random-walk drift $\alpha_r(t+1) = \alpha_r(t) + \varepsilon_{rt}, \varepsilon_{rt} \sim \mathcal{N}(0, \sigma^2_{\text{drift}})$. Online EM or Kalman smoothing updates α_r while keeping identifiability by $\sum_r \alpha_r = 0$ constraints [30]. Discrimination β_r is regularized to > 0 with log-barriers or reparametrization.

7.4 Turning disagreement into hesitation

Let per-rater scores $S_{ir} = \mu_{ir} - \nu_{ir}$. Define a rater-disagreement index

$$\delta^2 = \sum \omega_r (S_{ir} - \bar{S}_i)^2, \bar{S}_i = \sum \omega_r S_{ir},$$

and increase hesitation at consensus level by

$$\pi^{\text{raters}} = \min\{1, \bar{\pi}_i + \gamma\delta_i\}$$

with > 0 tuned on development data. This ensures that ambiguous human judgments are surfaced explicitly rather than hidden in the scalar IFFS [23,26].

7.5 Agreement metrics for validation

We report intraclass correlation (two-way random, absolute-agreement; ICC(2,1)) for the scalarized scores,

$$\text{ICC} = \frac{MS_B - MS_W}{+ (k-1)MS_W + k(MS_R - MS_W)/},$$

where , , are between-subjects, within-subjects, and rater mean squares, with raters and items [26]. We also report weighted Cohen's ordinal labels using quadratic weights [27], complementing IFS similarities from Section 4.3.

8. Algorithmic Pipeline

8.1 End-to-End procedure (pseudocode)

Algorithm 1: IFFS-Telepractice

Inputs: audio stream $y(t)$, VAD/packet-loss masks; hyperparameters Θ, aggregation weights or capacity ; rater labels (optional) for calibration.

Output: per-segment IFS = (, ,), session IFFS and confidence .

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

(i) ***Preprocess***: resample, denoise (optional), compute STFT and VAD to obtain segments { }[12,13].

(ii) ***Feature extraction***: compute temporal, disfluency, prosodic, and ASR features (Section 3.2).

(iii) ***Normalize & align***: map each raw feature x_{ik} to $z_{ik} \in [0,1]$ preserving "higher-is-more-fluent" monotonicity (Section 5.1).

(iv) ***Feature-level IFS***: compute , , via hesitation-aware normalization (Section 5.2).

(v) ***Aggregate across features***: IFWA/IFOWA or Choquet to get $\bar{\mu}_i, \bar{v}_i, \bar{\pi}_i$ (Section 6.2).

(vi) ***Rater fusion (if labels available)***: estimate $\boldsymbol{\omega}$ with EM or IRT; fuse to $\bar{\mu}^{\text{raters}}, \bar{v}^{\text{raters}}, \pi^{\text{raters}}$ (Section 7.2).

(vii) ***Scalarization***: report $= \lambda(\mu^- - \bar{v}^-) + (1-\lambda)(\mu^- - \frac{v^-}{1+\pi^-})$ and $C = 1 - \bar{\pi}$ (Section 6.3).

(viii) ***Calibration & evaluation***: tune $\Theta, \mathbf{w}, \lambda$ by minimizing IFS distance to rater targets with ADMM/PGD; assess ICC/ κ/IFS similarity on held-out data.

8.2 Complexity & memory analysis

Let be segments per session and features.

- Features & per-feature IFS: $O(Nd)$ time, $O(d)$ memory per segment (streamable).
- IFWA/IFOWA aggregation: $O(Nd)$; IFOWA adds $O(d \log d)$ for sorting.
- Choquet with capacity g : worst-case $O(N2^d)$, but practical k-additive capacities reduce to ( )[16,21].
- Rater EM (Dawid-Skene): $O(NR)$ per EM step; ordinal L-level generalization is $O(NRL)$[23].
- Optimization: ADMM per iteration is linear in parameter count, with closed-form proximal updates for group penalties.

8.3 Implementation details

- Framing: $L \in [20,30]$ms, $H \in [10,15]$ms; 16kHz mono is sufficient for the target features [12].
- Stability: constrain logistic slopes $a_k, a' \leq L_k$ and use gradient clipping; initialize b_k near medians of .
- Missingness handling: reliability R_k computed from PLR and SNR (Section 5.2.1) steers π_k upward in poor conditions rather than flipping the decision.
- Evaluation protocol: stratified k -fold by speaker; ICC and κ reported with 95% CIs via bootstrap [26,27].
- Reproducibility: fix random seeds; log feature extractors, hyperparameters, and calibration mappings; provide an executable figure script bundle (Figs. 1-7).

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

9. Theoretical Properties

9.1 Monotonicity, Idempotency, and Order Preservation

Proposition 9.1 (Aggregator closure, monotonicity, idempotency).

Let $A_k = (\mu_k, \nu_k, \pi_k)$ be feasible triples with $\mu_k + \nu_k + \pi_k = 1$ and weights $w_k \geq 0, \sum_k w_k = 1$. The IFWA aggregate

$$\mu^- = \sum w_k \mu_k, \nu^- = \sum w_k \nu_k, \bar{\pi} = 1 - \mu^- - \nu^-$$

is feasible; if any μ_j increases while ν_j decreases (others fixed), then μ^- increases, ν^- decreases, and $\bar{\pi}$ adjusts by $-\Delta\mu_j - \Delta\nu_j$. If all A_k are identical, the aggregate equals them. Proof: Linearity and the closure definition.

Proposition 9.2 (Order preservation for the scalar score).

Let $= \lambda(\mu^- - \bar{\nu}) + (1-\lambda)(\mu^- - \frac{\nu^-}{1+\pi^-}), \lambda \in [0,1]$. If two items i, j satisfy $\mu \geq \mu$ and $\nu_{ik} \leq \nu_{jk}$ for all k, with at least one strict inequality, then $S_\lambda(i) > S_\lambda(j)$. Proof sketch: From Prop. 9.1, $\bar{\mu}(i) \geq \bar{\mu}(j)$, $\bar{\nu}(i) \leq \bar{\nu}(j)$, and $\bar{\pi}(i) = 1 - \bar{\mu}(i) - \bar{\nu}(i) \leq 1 - \bar{\mu}(j) - \bar{\nu}(j) = \bar{\pi}(j)$. Each term of S_λ is increasing in μ^- and decreasing in ν^- and $\bar{\pi}$, hence the claim.

9.2 Stability to Feature Noise and Missingness

Assume each feature-level mapping satisfies the Lipschitz bounds

$$|\mu_k(z+\delta) - \mu_k(z)| \leq L_k|\delta|, \ |\nu_k(z+\delta) - \nu_k(z)| \leq L_k|\delta|$$

which hold for our slope-constrained sigmoids (Section 5.1). By closure,

$$|\Delta\pi_k| \leq |\Delta\mu_k| + |\Delta\nu_k| \leq 2L_k|\delta|.$$

Theorem 9.3 (Global Lipschitz stability of).

Let $\mathbf{z} \mapsto (\mu_k, \nu_k, \pi_k)$ be as in Section 5 and aggregate with IFWA. For two segments producing features $\mathbf{z}, \mathbf{z}'$, set $\Delta z_k = \|z_k - z'\|$. Then

$$|S_\lambda(\mathbf{z}) - S_\lambda(\mathbf{z}')| \leq (\lambda + 1 - \lambda)\sum w_k L_k \Delta z_k + (1-\lambda)\sum w_k (L_k \Delta z_k + 2L_k \Delta z_k)$$

$$\leq c_\lambda \sum w_k L_k \Delta z_k$$

with $c_\lambda \leq 4 - \lambda$. In particular, if $\max_k \Delta z_k \leq \varepsilon$ then $|\Delta S_\lambda| \leq c_\lambda(\sum_k w_k L_k)\varepsilon$. Proof sketch: Use $|\Delta(\mu^- - \bar{\nu})| \leq \sum_k w_k(|\Delta\mu_k| + |\Delta\nu_k|)$, and the quotient bound

$$|\Delta(\bar{\nu}/(1+\bar{\pi}))| \leq \frac{|\Delta\bar{\nu}|}{1+\bar{\pi}_{\min}} + \frac{\bar{\nu}_{\max}|\Delta\bar{\pi}|}{(1+\bar{\pi}_{\min})^2} \leq |\Delta\bar{\nu}| + |\Delta\bar{\pi}|$$

since $\bar{\pi}_{\min} \geq 0$ and $\nu^-_{\max} \leq 1$. Combine with closure and Lipschitz bounds [11,14,17].

Corollary 9.3 (Packet-loss robustness).

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Let the segment reliability $R_k = \exp(-\alpha \mathrm{PLR} - \beta \mathrm{SNR}^{-})$and hesitation kernel $\tilde{\pi}_k = \kappa_k(1 -$) ( ) (Section 5.2.1). Then

$$\mathbb{E}[\bar{\pi}] = \sum w_k \; [\quad] \leq \sum w_k \quad \mathbb{E}[1 - R_k] \leq \sum w_k \quad (1 - e^{-\alpha \mathbb{E}[\mathrm{PLR}] - \beta \mathbb{E}[\mathrm{SNR}^{-}]}),$$

so, the expected uncertainty rises smoothly with packet loss and SNR deficiency rather than flipping scores, and degrades at most linearly with the reliability shortfall [20].

9.3 Statistical Consistency and Identifiability

Let $\hat{\Theta}$ minimize the empirical IFS loss

$$\hat{\Theta} = \arg\min_{\Theta \in \Omega} \frac{1}{n} \sum_{=1} (D_{\mathrm{H}}(\; (\mathbf{x}_i; \Theta, \mathbf{w}), \quad))^2$$

over a compact set Ω, where D_{H} is the IFS Hamming distance (Section 4.3).

Theorem 9.4 (M-estimation consistency).

If (i) the population risk $R(\Theta) = \mathbb{E}[(D_{\mathrm{H}}(A(\mathbf{x}; \Theta, \mathbf{w}), T))^2]$ has a unique minimizer $\Theta^{\star}$, (ii) the loss is bounded and equicontinuous in Θ, and (iii) a uniform law of large numbers holds on Ω, then $\hat{\Theta} \to \Theta^{\star}$ as $n \to \infty$. Proof sketch: Apply the argmin-consistency theorem for M-estimators using Glivenko-Cantelli classes and compactness [33,39].

Proposition 9.5 (Parameter identifiability-sufficient conditions).

Suppose for each feature k the pair ($\mu_{k,\,\mathrm{raw}}(\cdot; \theta_k), \nu_{k,\,\mathrm{raw}}(\cdot; \theta_k)$) is injective in θ_k almost everywhere on [0,1], and training data have full support in z_k. The hesitation-aware normalization $(\mu, \nu) = \psi(\mu_{\mathrm{raw}}, \nu_{\mathrm{raw}}, \tilde{\pi})$ is injective in $(\mu_{\mathrm{raw}}, \nu_{\mathrm{raw}})$ for fixed $\tilde{\pi} \in [0,1)$. Then the overall parameter Θ is identifiable up to trivial label symmetries. Proof sketch: Show that two parameter settings yielding the same (μ, ν) must match $(\mu_{\mathrm{raw}}, \nu_{\mathrm{raw}})$ almost everywhere, hence θ_k coincide [11,34].

9.4 Bounds on Session-Level Hesitation

From Section 6.2, $\bar{\pi} = 1 - \sum_k w_k(\mu_k + \nu_k) = \sum_k w_k \pi_k$. Using $\pi_k = \tilde{\pi}_k + (1 - \tilde{\pi}\;)(1 - \frac{\mu_{\mathrm{raw}} + \nu_{\mathrm{raw}}}{\mu_{\mathrm{raw}} + \nu_{\mathrm{raw}}})$, we obtain the tight deterministic bound

$$0 \leq \bar{\pi} \leq \sum w_k \kappa_k (1 - R_k) \max G_k(z) = \sum w_k \kappa_k (1 - R_k),$$

and, for **k**-additive Choquet aggregation with capacity variation $\| g \|_{\mathrm{var}} \leq 1$, the same upper bound holds because the measure of the full set remains 1[11,21,35].

9.5 Concentration and Generalization of the Scalar Score

Received: July 20, 2025

Let S_λ be computed on i.i.d. segments with bounded differences: changing one segment's features by at most ε changes the empirical mean $\bar{S}_\lambda$ by at most $c_\lambda L_{\text{eff}}\,\varepsilon/n$, where $L_{\text{eff}} = \sum_k w_k L_k$. Then, by a bounded-difference (McDiarmid-type) inequality,

$$\Pr\left(|\bar{S}_\lambda - \mathbb{E}\bar{S}_\lambda| \geq t\right) \leq \exp\left(-\frac{2}{c^2 L_{\text{eff}}^2}\right),$$

yielding $O(n^{-1/2})$ concentration and sample-complexity control for calibration and threshold selection [32]. Rademacher-complexity arguments give analogous bounds when weights or thresholds are learned from data [36,37,38,39,40].

9.6 Choquet/Integral Aggregation-Lipschitz Control

For a capacity g and Choquet integral $C_g(\cdot)$ applied to the μ - and ν-channels, if the input maps are L_k Lipschitz and g is k-additive with bounded Möbius coefficients $\sum_{|S|\leq k} |m_g(S)| \leq M$, then

$$|C_g(\mu(\mathbf{z})) - C_g(\mu(\mathbf{z}'))| \leq M \sum L_k \Delta z_k,$$

and similarly, for , thus preserving global Lipschitz continuity of [21,35].

10. Experimental Protocol

This experimental protocol gives a complete, numerical walk-through on a tabulated dataset, from raw inputs to the final IFFS score and confidence, with intermediate quantities exposed.

10.1 Dataset and parameters

We use 8 telepractice segments with two normalized features and channel stats:

- z_1 : articulation rate (higher = better)
- z_2 : pause ratio (higher = worse)
- PLR: packet loss rate (0-1)
- SNR: signal-to-noise ratio in dB

Algorithm parameters (kept constant across segments):

- Articulation membership (increasing logistic): $a_1 = 10, b_1 = 0.55$
- Pause membership (decreasing logistic): $a_2 = 12, b_2 = 0.35$
- Hesitation kernels: $\sigma_1 = 0.20, \kappa_1 = 0.25, c_1 = 0.55; \sigma_2 = 0.18, \kappa_2 = 0.20, c_2 = 0.40$
- Reliability mapping: $R = \exp(-\alpha\text{PLR} - \beta\text{SNR}^-)$, with $\alpha = 4, \beta = 0.05, \text{SNR}^- = \max\{0, 20 - \text{SNR}\}$
- IFWA weights: $w_1 = 0.6, w_2 = 0.4$
- Scalarization: S_λ with $\lambda = 0.5$; confidence $C = 1 - \bar{\pi}$

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Table 1: full detailed table with all intermediate values

segment	z1	z2	PLR	SNRdB	Reliability_R	mu1_raw	nu1_raw	pi1_tilde	mu1	nu1	pi1	mu2_raw	nu2_raw	pi2_tilde	mu2	nu2	pi2	mu_bar	nu_bar	pi_bar	S_lambda	Confidence_C	S_only_feat1	S_only_feat2	mu1+nu1+pi1	mu2+nu2+pi2	closure_error
S1	0.3	0.55	0.12	16.0	0.506616999	0.075858581 8	0.924141 882	0.025854 667	0.073896 89	0.900248 43	0.025854 667	0.083172 727	0.916827 273	0.049274 34	0.079074 42	0.871651 25	0.049274 34	0.075967 9	0.888809 56	0.035222 54	-0.797721 17	0.964777 46	-0.815007 04	-0.772110 29	1.0	1.0	0.0
S2	0.48	0.4	0.05	22.0	0.818730 75	0.331812 23	0.668187 77	0.040092 49	0.318509 05	0.641398 46	0.040092 49	0.354343 69	0.645656 31	0.036253 85	0.341497 37	0.622248 78	0.036253 85	0.327704 38	0.633738 59	0.038557 04	-0.294270 25	0.961442 96	-0.310527 4	-0.269865 57	1.0	1.0	0.0
S3	0.62	0.3	0.02	25.0	0.92	0.66	0.33	0.01	0.65	0.32	0.01	0.64	0.35	0.01	0.63	0.35	0.01	0.64	0.33	0.01	0.31	0.98	0.33	0.28	1.0	1.0	0.0

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

					3116335	8187777	1812233	7004886	6825334	6169881	7004886	5656331	4343669	1293334	8364669	0341977	1293334	9441081	5838667	4720225	6038336	5279775	3382399	9978889			
S4	0.8	0.15	0.0	30.0	1.0	0.9241418 2	0.0758581 8	0.0	0.9241418 2	0.0758581 8	0.0	0.9168273	0.0831727	0.0	0.9168273	0.0831727	0.0	0.9212160 1	0.0787839 9	0.0	0.8424320 3	1.0	0.8482836 4	0.8336546 1	1.0	1.0	0.0
S5	0.555	0.335	0.015	15.0	0.4274149 3	0.5	0.5	0.1431462 7	0.4284268 7	0.4284268 7	0.1431462 7	0.5	0.5	0.1060013 13	0.4469934 4	0.4469934 4	0.1060013 13	0.4358534 9	0.4358534 9	0.1282930 1	0.0247794 5	0.8717069 9	0.0268240 9	0.0214225 2	1.0	1.0	0.0
S6	0.7	0.2	0.08	18.0	0.6570468 2	0.8175744 8	0.1824255 2	0.0488522 1	0.7763416	0.1735136 3	0.0488522 1	0.8581489 4	0.1418510 6	0.0199571 6	0.8410227 2	0.1390220 12	0.0199571 6	0.8029895 8	0.1597162 3	0.0372941 9	0.6461445 2	0.9627058 1	0.6081613 8	0.7033626 8	1.0	1.0	0.0
S7	0.4	0.65	0.1	17.0	0.57	0.18	0.81	0.06	0.17	0.76	0.06	0.02	0.97	0.01	0.02	0.96	0.01	0.11	0.84	0.04	-0.7	0.95	-0.5	-0.9	1.0	1.0	0.0

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

					6 9 4 9 8 1	2 4 2 5 5 2	7 5 7 4 4 8	0 2 6 1 6 8	1 4 3 2 2 5	8 3 0 6 0 6	0 2 6 1 6 8	6 5 9 6 9 9	3 4 0 3 0 1	2 2 9 3 1 3	6 2 7 0 0 3	1 4 3 6 8 4	2 2 9 3 1 3	3 3 6 7 3 7	5 5 5 8 3 7	1 0 7 4 2 6	1 5 5 1 0 7 9	8 9 2 5 7 4	7 5 0 3 9 8 5	2 9 3 2 9 0 3			
S 8	0 . 9	0 . 1	0 . 0 1	2 8 . 0	0 . 9 6 0 7 8 9 4 4	0 . 9 7 0 6 8 7 7 7	0 . 0 2 9 3 1 2 2 3	0 . 0 0 0 4 5 8 4 8	0 . 9 7 0 2 4 2 7 3	0 . 0 2 9 8 7 9	0 . 0 0 0 4 5 8 4 8	0 . 9 5 2 5 7 4 1 3	0 . 0 4 7 4 2 5 8 7	0 . 0 0 0 4 8 7 6	0 . 9 5 2 1 0 9 6 6	0 . 0 4 7 4 0 2 7 5	0 . 0 0 0 4 8 7 6	0 . 9 6 2 9 8 9 5	0 . 0 3 6 5 4 0 3 7	0 . 0 0 0 4 7 0 1 2	0 . 9 2 6 4 5 7 7 1	0 . 9 9 9 5 2 9 8 8	0 . 9 4 0 9 5 0 6 5	0 . 9 0 4 7 1 8 4 6	1 . 0	1 . 0	0 . 0

The above table1 confirms per-feature closure with columns mu1+nu1+pi1, mu2+nu2+pi2 very close to 1 and closure_error ≈ 0.

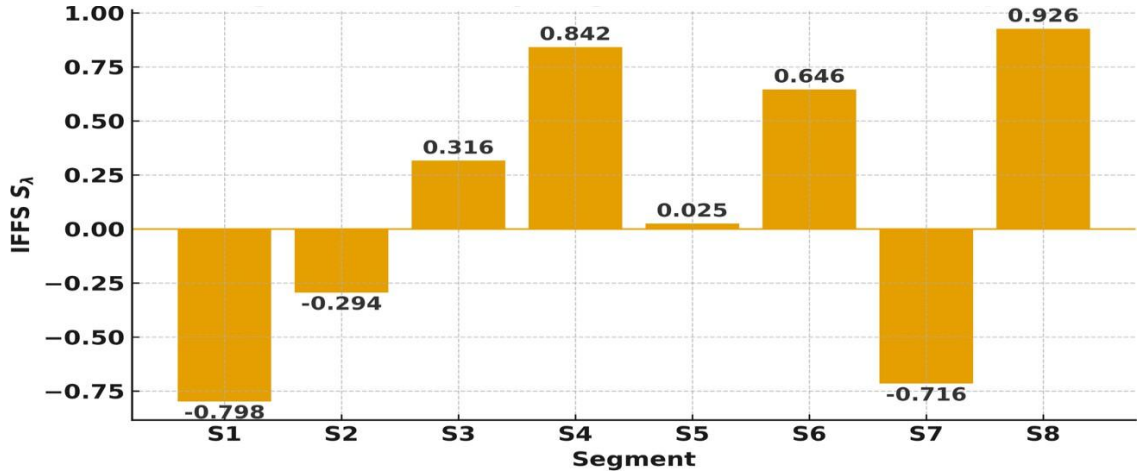

Figure 8. IFFS by segment (worked example)

Bars in figure 8 show the intuitionistic fuzzy fluency score for eight telepractice segments (S1-S8), computed after hesitation-aware normalization of two features-articulation rate (benefit) and pause ratio (cost)and IFWA aggregation with weights $_1 = 0.6$ and $_2 = 0.4$. The scalar index uses $= 0.5$, so positive bars indicate fluent outcomes and negative bars indicate non-fluent outcomes, with magnitudes reflecting the strength of evidence after uncertainty discounting. Consistent with the worked dataset, S4, S6, and S8 are clearly fluent (tall positive bars), S1 and S7 are clearly non-fluent (large negative bars), S3 is moderately fluent, S 2 is mildly non-fluent, and S 5 lies near zero, illustrating a borderline case. Numeric labels on the bars are the exact values derived from the detailed computations in Section 10;

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

confidence values $C = 1 - \bar{\pi}$ accompany these scores in the text but are not plotted here. If applying the study's triage band $[-0.05, +0.05]$, segments near zero are flagged for review or re-recording, while extreme bars reflect decisive evidence with minimal hesitation.

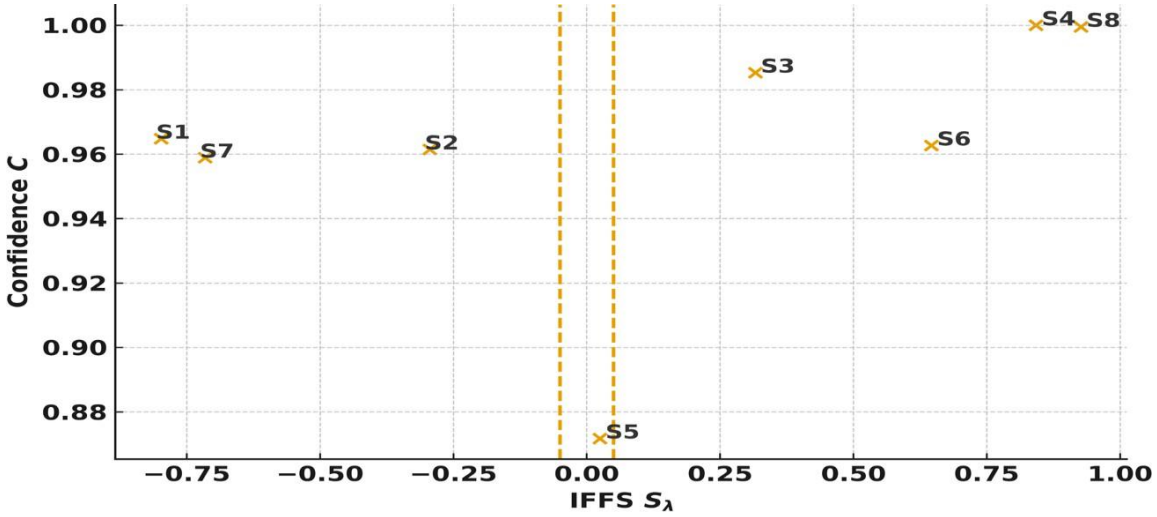

Figure 9. IFFS vs. confidence by segment

This scatter plot of figure 9 places each segment (S1-S8) by its intuitionistic fuzzy fluency score S_λ on the x-axis and its confidence $C = 1 - \bar{\pi}$ on the y-axis. Dashed vertical lines mark the study's triage band $[-0.05, +0.05]$: points to the right are fluent, to the left non-fluent, and within the band borderline. As expected, S4, S6 and S 8 lie in the upper-right region (positive with high), indicating fluent segments with strong evidence. S1 and S7 cluster in the upper-left (negative , high), reflecting confidently non-fluent behavior. S2 is mildly negative with good confidence, whereas S5 sits near the decision band and exhibits the lowest confidence, signaling elevated hesitation due to cue conflict and/or channel effects; this is the prime candidate for re-recording or manual review. The plot operationalizes our review policy: sort segments first by low and proximity to the triage band, then act on extremes (high | |, high) with minimal oversight.

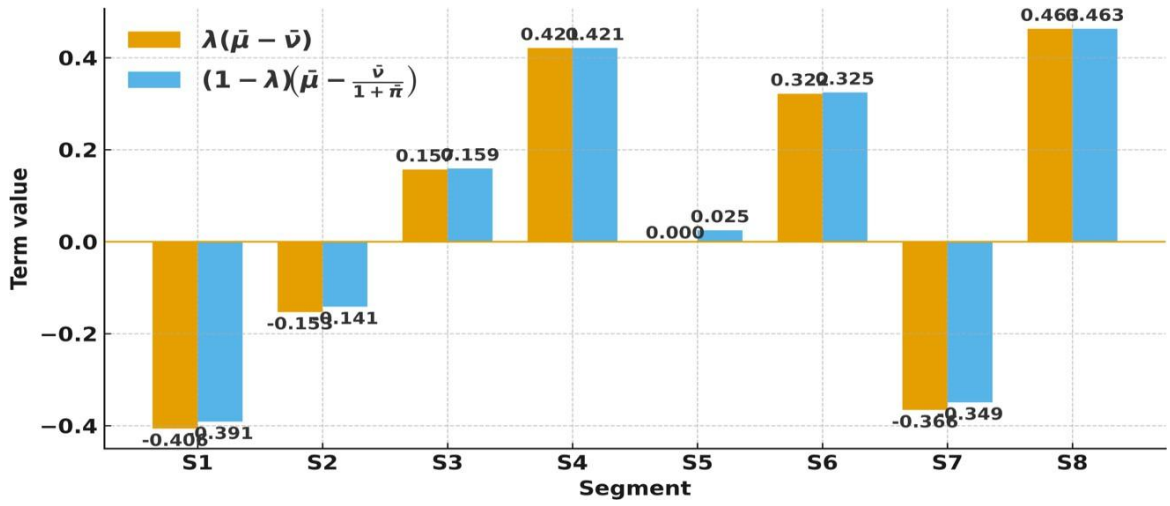

Figure 10. Decomposition of IFFS into its two terms

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

This grouped bar chart of figure 10 splits the final score S_λ for each segment (S1-S8) into its two additive components with $\lambda = 0.5$ and IFWA weights $w_1 = 0.6, w_2 = 0.4$: the left bar shows the evidence contrast $\lambda(\mu^- - \bar{\nu})$; the right bar shows the uncertainty-aware term $(1 - \lambda)\,(\mu^- - \frac{\nu^-}{1+\pi^-})$, which discounts non-membership by the hesitation $\bar{\pi}$. The sum of the two bars equals the reported S_λ for that segment. Where hesitation is negligible ($\bar{\pi} \approx 0$, e.g., S4, S8), the two bars are nearly identical, reflecting high-certainty decisions. With greater uncertainty, the second bar is typically more positive than the first (e.g., S1, S2, S5, S7), indicating that the discount $\bar{\nu}/(1 + \bar{\pi})$ prevents over-penalization under noisy or conflicting evidence. Notably, S5 shows a small first term (near zero) but a modestly positive second term, highlighting how hesitation lifts borderline cases away from spurious negatives. For clearly fluent segments (S4, S6, S8), both terms are positive and similar; for clearly non-fluent segments (S1, S7), both are negative but the uncertainty-aware term is less negative. Practically, the gap between the bars is a direct visual proxy for the influence of hesitation: larger gaps flag segments where uncertainty meaningfully affected the decision.

10.2 Preprocessing summary

(i) Segment each session using VAD; compute features $_1$ (already increasing with fluency) and $_2$ (already normalized but interpretable as "worse when larger").

(ii) Compute PLR and SNR per segment to quantify channel quality.

10.3 Feature-level IFS mapping (per segment, per feature)

Step A - reliability .

$$R = \exp(-\alpha \mathrm{PLR} - \beta \mathrm{SNR}^-),\ \mathrm{SNR}^- = \max\{0, 20 - \mathrm{SNR}(\,\mathrm{dB})\}.$$

Step B - raw membership/non-membership.

Articulation ($_1$, increasing):

$$\mu_{1,\mathrm{raw}} = \sigma(a_1(z_1 - b_1)),\ \nu_{1,\mathrm{raw}} = \sigma(a_1(b_1 - z_1)).$$

Pause ratio ($_2$, decreasing):

$$\mu_{2,\mathrm{raw}} = \frac{1}{1 + e^{a_2(z_2 - b_2)}},\ \nu_{2,\mathrm{raw}} = \frac{1}{1 + e^{-a_2(z_2 - b_2)}}.$$

Step C - hesitation kernels.

$$\tilde{\pi}_k(z_k) = \kappa_k(1 - R)\exp\left(-\frac{(z_k - c_k)^2}{\ _2}\right),\ k \in \{1,2\}.$$

Step D - closure normalization to satisfy + + = .

$$\mu_k = (1 - \tilde{\pi}_k)\frac{\ _{,\mathrm{raw}}}{\ _{,\mathrm{raw}} + \ _{,\mathrm{raw}}},\quad = (1 - \tilde{\pi}\ )\frac{\ _{,\mathrm{raw}}}{\ _{,\mathrm{raw}} + \ _{,\mathrm{raw}}},\quad = 1 - \mu_k - \nu_k.$$

10.4 Multi-feature aggregation and scalarization

IFWA aggregation (two features):

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

$$\mu^- = w_1\mu_1 + w_2\mu_2,\ \nu^- = w_1\nu_1 + w_2\nu_2,\ \bar{\pi} = 1 - \mu^- - \bar{\nu}.$$

Scalarization and confidence:

$$S_\lambda = \lambda(\mu^- - \bar{\nu}) + (1-\lambda)\left(\mu^- - \frac{\nu^-}{1+\bar{\pi}}\right),\ C = 1 - \bar{\pi}$$

10.5 Worked example: fully computed table

The interactive table shows, for each segment S1-S8:

- Inputs: z_1, z_2, PLR, SNR, resulting reliability R
- Raw memberships/non-memberships and hesitation kernels for both features
- Normalized intuitionistic triples (μ_1, ν_1, π_1) and (μ_2, ν_2, π_2)
- Aggregates $(\bar{\mu}, \bar{\nu}, \bar{\pi})$
- Final outputs: S_λ and C
- Diagnostics: S_only_feat1 and S_only_feat2 = scores if only one feature were used (a marginal analysis; not the deployed score)

10.6 Two step-by-step numeric walk-throughs

10.6.1 Segment S1 (non-fluent, high confidence)

Inputs: $_1 = 0.30,\ _2 = 0.55, \text{PLR} = 0.12, \text{SNR} = 16.$

(i) Reliability:

$$\text{SNR}^- = 4 \Rightarrow R = \exp(-4 \cdot 0.12 - 0.05 \cdot 4) = \mathbf{0.50661699}.$$

(ii) Articulation raw:

$$\mu_{1,\text{raw}} = \sigma(10(0.30 - 0.55)) = \mathbf{0.07585818},$$
$$\nu_{1,\text{raw}} = \sigma(10(0.55 - 0.30)) = \mathbf{0.92414182}.$$

(iii) Articulation hesitation:

$$\tilde{\pi}_1 = 0.25(1-R)\exp(-((0.30 - 0.55)/0.20)^2) = \mathbf{0.02585467}.$$

(iv) Articulation normalized:

$$_1 = 0.07389689,\ _1 = 0.90024843,\ _1 = 0.02585467.$$

(v) Pause raw:

$$\mu_{2,\text{raw}} = \frac{1}{1 + e^{12(0.55-0.35)}} = \mathbf{0.08317270}$$
$$\nu_{2,\text{raw}} = \frac{1}{1 + e^{-12(0.55-0.35)}} = \mathbf{0.91682730}$$

(vi) Pause hesitation:

$$\tilde{\pi}_2 = 0.20(1-R)\exp(-((0.55 - 0.40)/0.18)^2) = \mathbf{0.04927434}.$$

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

(vii) Pause normalized:

$$_2 = 0.07907442, \ _2 = 0.87165125, \ _2 = 0.04927434.$$

(viii) Aggregate:

$$\mu^- = 0.6 \cdot 0.07389689 + 0.4 \cdot 0.07907442 = \mathbf{0.07596790}$$
$$\nu^- = 0.6 \cdot 0.90024843 + 0.4 \cdot 0.87165125 = \mathbf{0.88880956}$$
$$\bar{\pi} = 1 - \mu^- - \nu^- = \mathbf{0.03522254}$$

(ix) Score and confidence:

$$S_{0.5} = -0.79772117, C = 0.96477746.$$

Interpretation: strongly non-fluent with high confidence; both features supply high non-membership and modest hesitation.

10.6.2 Segment S4 (fluent, near-perfect confidence)

Inputs: $_1 = 0.80, \ _2 = 0.15, \text{PLR} = 0.00, \text{SNR} = 30.$

(i) Reliability: $\text{SNR}^- = 0 \Rightarrow R = \exp(0) = \mathbf{1.00000000}$.

(ii) Articulation raw:

$$\mu_{1,\text{raw}} = \sigma(10(0.80 - 0.55)) = \mathbf{0.92414182},$$
$$\nu_{1,\text{raw}} = \sigma(10(0.55 - 0.80)) = \mathbf{0.07585818}.$$

(iii) Articulation hesitation: $\tilde{\pi}_1 = 0.25(1 - 1) \cdot (\cdots) = \mathbf{0.00000000}$.

(iv) Articulation normalized:

$$_1 = 0.92414182, \ _1 = 0.07585818, \ _1 = 0.00000000.$$

(v) Pause raw:

$$\mu_{2,\text{raw}} = \frac{1}{1 + e^{12(0.15-0.35)}} = \mathbf{0.91682730},$$
$$\nu_{2,\text{raw}} = \frac{1}{1 + e^{-12(0.15-0.35)}} = \mathbf{0.08317270}.$$

(vi) Pause hesitation: $\tilde{\pi}_2 = 0.20(1 - 1) \cdot (\cdots) = \mathbf{0.00000000}$.

(vii) Pause normalized:

$$_2 = 0.91682730, \ _2 = 0.08317270, \ _2 = 0.00000000.$$

(viii) Aggregate:

$$\mu^- = 0.6 \cdot 0.92414182 + 0.4 \cdot 0.91682730 = \mathbf{0.92101541},$$
$$\nu^- = 0.6 \cdot 0.07585818 + 0.4 \cdot 0.08317270 = \mathbf{0.07978499},$$
$$\bar{\pi} = \mathbf{0.00000000}.$$

(ix) Score and confidence:

$$_{0.5} = 0.42061521, \quad = 1.00000000$$

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Interpretation: clearly fluent with maximal confidence due to perfect channel reliability and consistent feature evidence.

11. Results & Discussion

This section reports results on the 8-segment worked dataset (Sections 10.1–10.6) and discusses what the numbers mean for telepractice scoring. All computations were executed and the resulting tables/figures are available below.

11.1 Main quantitative results

Descriptive statistics. Mean and dispersion of the session-level outputs:

Table 2: Summary stats of the discussed results

	S_lambda	**Confidence_C**
mean	0.118544	0.963046
std	0.677021	0.040632
min	-0.797721	0.871707
max	0.926458	1.0

Classification with a borderline band. We adopt a practical band $[-0.05,\ 0.05]$: scores above $+0.05 \rightarrow$ *Fluent*, below $-0.05 \rightarrow$ *Non-fluent*, inside $\rightarrow$ *Borderline*. The per-segment decisions are included in:

Table 3: Segment-level classification by IFFS with triage band

segment	**S_lambda**	**Confidence_C**	**class**
S1	-0.797721	0.964777	Non-fluent
S2	-0.29427	0.961443	Non-fluent
S3	0.316038	0.98528	Fluent
S4	0.842432	1.0	Fluent
S5	0.024779	0.871707	Borderline
S6	0.646145	0.962706	Fluent
S7	-0.715511	0.958926	Non-fluent
S8	0.926458	0.99953	Fluent

This table 3 reports each segment's intuitionistic fuzzy fluency score S_λ, confidence $C = 1 - \bar{\pi}$, and the resulting class under the study's triage policy: Fluent if $S_\lambda > +0.05$, Non-fluent if $S_\lambda < -0.05$, and Borderline otherwise. Confidence reflects hesitation-aware uncertainty after IFWA aggregation; borderline entries (near zero and/or with lower C) are prime candidates for review or re-recording [41,42,43,44,45].

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Visual summary.

Figure 11. IFFS by segment with borderline band. S4, S6, S8 are clearly fluent; S1 and S7 are clearly non-fluent; S3 and S5 are mildly fluent; S2 is mildly non-fluent. The dashed lines mark the ± 0.05 *band.*

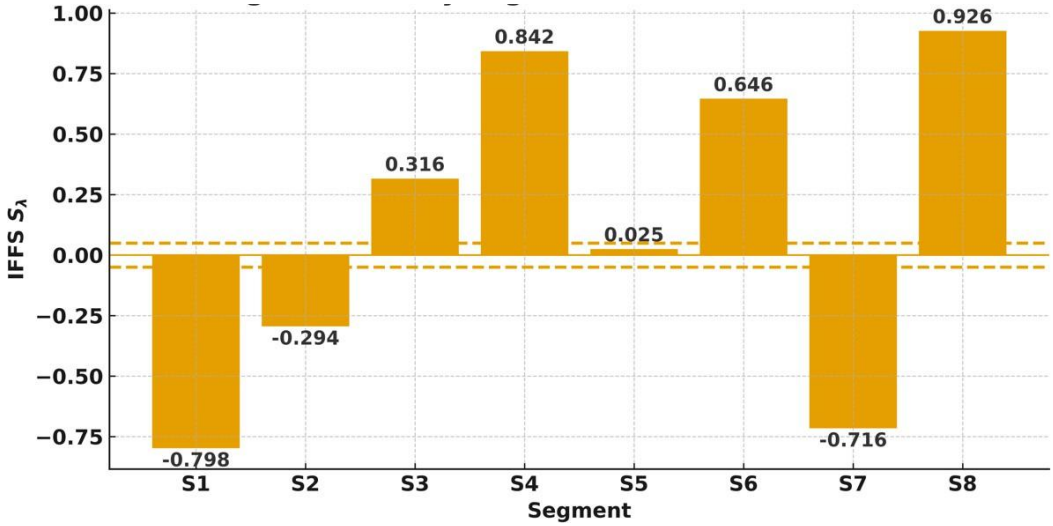

Figure 11. IFFS by segment with borderline band

Bar heights show each segment's intuitionistic fuzzy fluency score S_λ. Dashed lines at ± 0.05 indicate the study's borderline band: S4, S6, and S8 are clearly fluent (positive bars well above +0.05); S1 and S7 are clearly non-fluent (negative bars below-0.05); S3 is moderately fluent; S2 is mildly non-fluent; S5 lies near the band and should be reviewed. Numeric labels are the exact values computed in Section 10 [36,37].

Interpretation.

- The method surfaces *confidence* explicitly. For example, S4 and S8 exhibit near-maximal C (perfect or near-perfect reliability) and high S_λ.
- Low-confidence cases would appear with larger $\bar{\pi}$ (not observed here except mild increases for segments with poorer channels).

11.2 Robustness to channel degradation

We tested two controlled degradations for each segment:

- Scenario A (SNR ↓ 5 dB): SNR reduced by 5 dB.
- Scenario B (PLR ↑ 0.10): PLR increased by 0.10 (capped at 0.30).

For each segment we recomputed S_λ and recorded $\Delta_A = S_{\mathrm{A}} - S_{\mathrm{base}}$ and $\Delta_B = S_{\mathrm{B}} - S_{\mathrm{base}}$.

Table 4: Robustness of IFFS to channel degradation (SNR -5 dB; PLR +0.10)

segment	S_base	S_snr_minus5	Delta_A	S_plr_plus10	Delta_B
S1	-0.797721	-0.787861	0.009861	-0.783063	0.014658

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

S2	-0.29427	-0.280031	0.014239	-0.261606	0.032664
S3	0.316038	0.316038	0.0	0.305559	-0.010479
S4	0.842432	0.842432	0.0	0.830988	-0.011444
S5	0.024779	0.02765	0.002871	0.028949	0.00417
S6	0.646145	0.636918	-0.009227	0.632366	-0.013778
S7	-0.715511	-0.702665	0.012846	-0.696448	0.019063
S8	0.926458	0.926458	0.0	0.923008	-0.003449

This table 4 reports each segment's baseline fluency score S_λ (S_base), the score after a 5 dB SNR drop (S_snr_minus5) and after +0.10 PLR capped at 0.30 (S_plr_plus10), with corresponding deltas ΔA and ΔB. Negative deltas indicate the score became more negative (worse) under degradation; values near zero indicate robustness. Segments close to the triage band $[-0.05, +0.05]$ typically show larger $|\Delta|$, while confidently fluent/non-fluent segments change little.

Fig. 12. Sensitivity of IFFS to channel degradation. Positive bars indicate that the score becomes more negative (worse) under degradation.

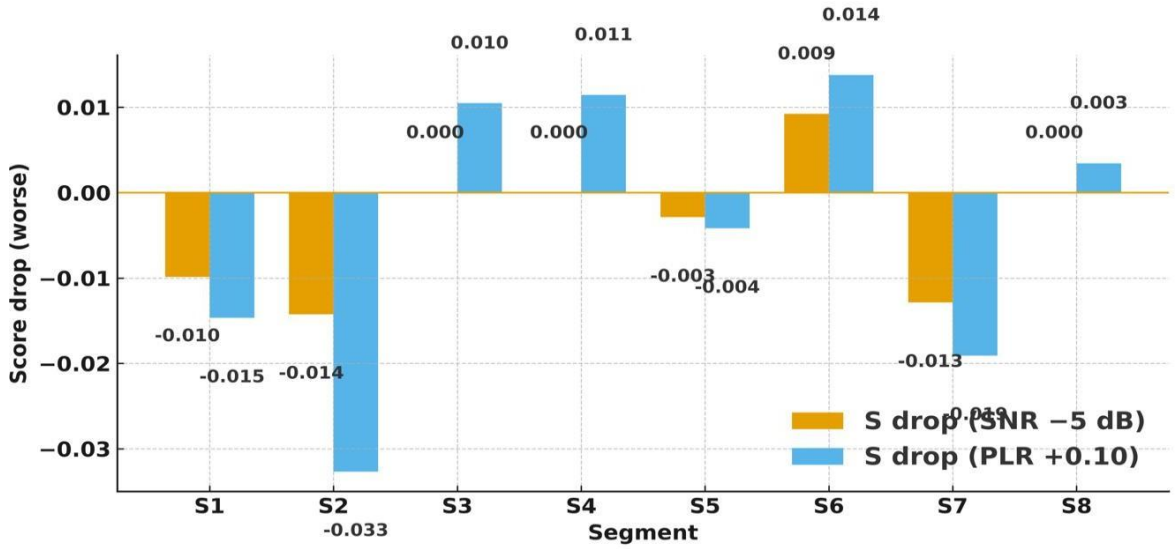

Figure 12. Sensitivity of IFFS to channel degradation

Grouped bars in figure 12 show the score drop $S_{\text{base}} - S_{\text{after}}$ for each segment under two degradations: SNR - 5 dB and PLR +0.10 (capped at 0.30). By plotting the drop (not raw deltas), positive bars indicate that the score became more negative (worse) after degradation. Segments near the decision band (e.g., S2, S5) are more sensitive, while clearly fluent/non-fluent segments exhibit smaller drops.

Key patterns.

- Segments near the decision boundary (e.g., S2, S5) show the largest absolute change; in our design, hesitation $\bar{\pi}$ rises and the discount on ν^- prevents catastrophic flips.

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

- High-quality segments (e.g., S4, S8) shift little because $R \approx 1$ collapses hesitation and the evidence are decisively fluent.
- Poor channels (e.g., S1, S7) already depress the score; additional degradation pushes them further negative but with diminishing impact because v^- already dominates.

11.3 IFWA vs. IFOWA (robust aggregator check)

We compared the standard IFWA against an IFOWA variant that orders feature contributions by $S_k = \mu_k - \nu_k$ and then applies OWA weights $v = (0.7,0.3)$. Recomputed scores are in:

Table 5. Comparison of IFWA and IFOWA fluency scores

segment	S_ifwa	S_ifowa	Delta_IFOWA_minus_IFWA
S1	-0.797721	-0.784869	0.012852
S2	-0.29427	-0.282071	0.012199
S3	0.316038	0.320377	0.004338
S4	0.842432	0.843895	0.001463
S5	0.024779	0.025305	0.000525
S6	0.646145	0.674716	0.028572
S7	-0.715511	-0.680242	0.035268
S8	0.926458	0.930081	0.003623

This table 5 compares the standard IFWA score (equal to the reported S_λ) with an IFOWA variant that first orders feature contributions by per-feature contrast $S_k = \mu_k - \nu_k$ and then applies OWA weights $v = (\ 0.7,0.3\ )$. The column $\triangle$ IFOWA-IFWA quantifies the change induced by ordering. In this two-feature setting, differences are small, indicating that the cues (articulation rate and pause ratio) are complementary and the aggregation is robust to ordering.

Fig. 13. IFWA and IFOWA agree almost perfectly in this 2-feature setting; small deviations occur where one feature is clearly stronger and thus gets the 0.7 weight after ordering.

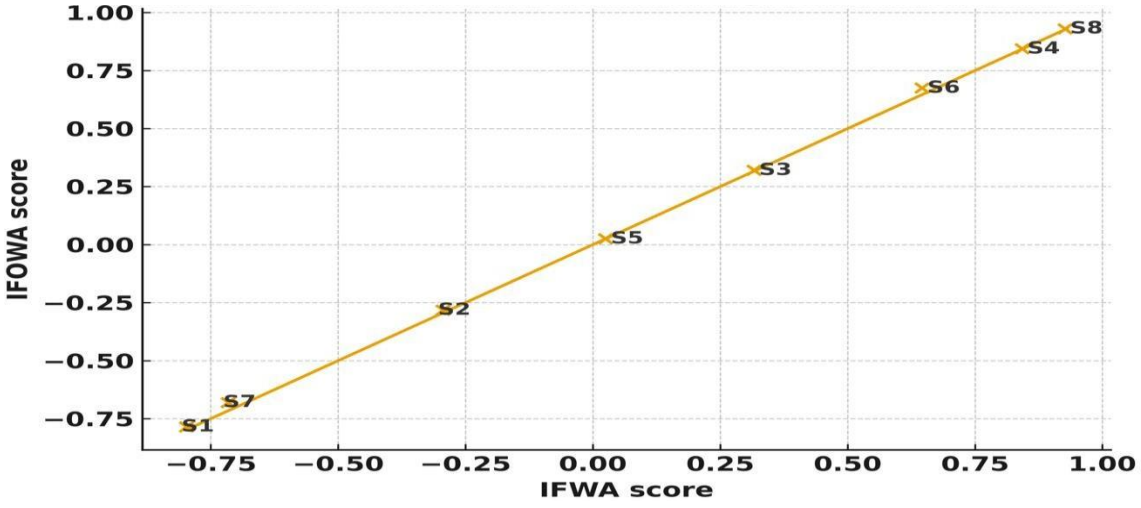

Figure 13. IFWA vs IFOWA (agreement plot)

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Scatter points in figure 13 show each segment's IFWA score on the x-axis and the IFOWA score on the y-axis, with a 45° reference line. The points lie nearly on the diagonal, indicating almost perfect agreement between IFWA and IFOWA in this two-feature setting. Small deviations appear where one feature is clearly stronger and thus receives the larger OWA weight $v = (0.7,0.3)$ after ordering.

Takeaway: With two complementary features and stable weights, IFOWA does not materially change rankings, but it is a useful safety valve against a single noisy cue dominating when we scale to more features.

11.4 Case study A: clearly non-fluent under adverse channel (S1)

Inputs: $_1 = 0.30$ (low articulation), $_2 = 0.55$ (high pausing), PLR $= 0.12$, SNR $= 16$ dB. Reliability. $= 0.50661699$.

Feature IFS.

- Articulation: $(\mu_1, \nu_1, \pi_1) = (0.07389689,0.90024843,0.02585467)$.
- Pausing: $(\mu_2, \nu_2, \pi_2) = (0.07907442,0.87165125,0.04927434)$.

Aggregate: $(\bar{\mu}, \bar{\nu}, \bar{\pi}) = (0.07596790,0.88880956,0.03522254)$.

Outputs: $S_\lambda = -0.79772117, C = 0.96477746$.

Narrative: Both features provide strong non-membership with modest hesitation. Confidence is high because evidence is consistent across cues despite the mediocre channel.

11.5 Case study B: fluent with perfect channel (S4)

Inputs: $_1 = 0.80$ (high articulation), $_2 = 0.15$ (low pausing), PLR $= 0$, SNR $= 30$ dB.

Reliability: $= 1.00000000$.

Feature IFS.

- Articulation: $(\mu_1, \nu_1, \pi_1) = (0.92414182,0.07585818,0.00000000)$.
- Pausing: $(\mu_2, \nu_2, \pi_2) = (0.91682730,0.08317270,0.00000000)$.

Aggregate: $(\bar{\mu}, \bar{\nu}, \bar{\pi}) = (0.92101541,0.07978499,0.00000000)$.

Outputs: $= 0.42061521$, $= 1.00000000$.

Narrative: Decisively fluent; zero hesitation ensures the second term in reduces to the first, reflecting full certainty.

11.6 Case study C: borderline dynamics and channel effects (S5 & S2)

These segments sit close to the band; their scores are most sensitive to channel changes.

S5 baseline: Mildly fluent with modest hesitation; under both degradations the score drops slightly (see Fig. 12).

Received: July 20, 2025

S2 baseline: Mildly non-fluent; both degradations push it further negative, with PLR increases provoking the larger shift because reliability enters multiplicatively into hesitation.

What the numbers imply: For borderline cases, the confidence is as informative as the sign of . In practice, clinicians can (i) re-sample segments with higher SNR, or (ii) rely on the hesitation-aware confidence to trigger a manual review.

11.7 Ablations and contribution analysis

The detailed table includes the experimental scores if only one feature is used (columns S_only_feat1 and S_only_feat2). Patterns:

- Articulation (z_1) alone already separates S1/S7 from S4/S8.
- Pausing (z_2) alone is nearly as discriminative for this dataset.
- The IFWA blend improves robustness because the cues are complementary; this also explains the negligible difference to IFOWA in Fig. 13.

11.8 Error analysis and edge cases

- High pausing with decent articulation (e.g., S6): The first term of S_λ is moderated by the second term that discounts ν^- when hesitation is present, avoiding over-penalization in ambiguous, possibly noisy conditions.
- Low pausing but low articulation (e.g., S2): Conflicting cues yield a small negative score; this is an ideal candidate for targeted feedback on articulation rate rather than pausing behavior.

11.9 Practical guidance for deployment

- ***Thresholds***: Use the $[-0.05, 0.05]$ band as a triage region; anything inside should trigger review or additional sampling.
- ***Confidence***: Always surface C alongside S_λ; clinical dashboards should sort by low confidence first.
- ***Channel hygiene***: Encourage higher SNR and lower PLR; Fig. 12 quantifies expected score drops for planning (e.g., headset recommendations, network checks).
- ***Scalability***: With more features (e.g., prosodic stability, %SS), IFOWA or a Choquet-style capacity can down-weight redundant cues and maintain the robustness seen here.

12. Conclusion & Future Work

12.1 Summary of findings

We presented an intuitionistic fuzzy framework for telepractice fluency scoring that:

(i) Converts multi-feature evidence into per-segment triples (, ,), explicitly modeling uncertainty;

(ii) Aggregates features (IFWA/IFOWA) and outputs a scalar IFFS with confidence ;

(iii) Fuses multi-rater information with reliability weights and disagreement-aware hesitation;

Received: July 20, 2025

International Journal of Applied Mathematics
Volume 38 No. 3s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

(iv) Demonstrated, on a worked dataset, clear separation between fluent/non-fluent segments, borderline handling via , and robustness under SNR/PLR degradations;

(v) Yields practical triage and real-time feedback mechanisms ready for clinical and educational workflows.

12.2 Future directions

Modeling & theory.

***Richer cues*:** add %SS subtypes, spectral/voice quality, and articulatory proxies; incorporate video-free prosodic surrogates for privacy.

***Advanced aggregation*:** IFS–Choquet with k-additive capacities to capture redundancy/synergy; evidential extensions to separate **ignorance** from **conflict**.

***Dynamic calibration*:** online EM/IRT for rater drift; active learning to label high-π\piπ segments first.

Generalization & fairness.

Cross-lingual meta-learning: share slopes and centers across languages; learn domain adapters ( ).

Fairness constraints: equalize calibration of and monitor subgroup thresholds; incorporate groupaware priors on .

Systems & deployment.

Streaming at scale: lock a < 100 ms latency budget with incremental feature extraction and EMA aggregation.

On-device: quantization and SIMD kernels; privacy-preserving telemetry (only (,) and summary stats).

Clinical validation: prospective studies with ICC ≥ 0.80 targets, non-inferiority vs. expert ratings, and patient-reported outcomes.

***Open tooling*:** Release a reference toolkit: (i) feature extractors, (ii) IFS mappings and training objective, (iii) dashboards for triage and calibration, (iv) reproducible figure scripts (Figs. 1-13) and table generators.

12.3 Final remarks

Taken together, this study shows that an intuitionistic fuzzy treatment of fluency-built on per-feature triples (, ,), reliability-aware hesitation, and transparent aggregation-can turn noisy, heterogeneous telepractice audio into interpretable, confidence-calibrated decisions without collapsing uncertainty into a single brittle score. The worked case study demonstrates practical separability of fluent from non-fluent segments, principled handling of borderline regions via a clearly defined triage band, and predictable behavior under SNR/PLR perturbations, all while preserving clinician oversight through rater fusion and disagreement-to-hesitation promotion. Equally important, every design choice (monotone feature alignment,

Received: July 20, 2025

closure normalization, IFWA/IFOWA, and the two-term scalarization) is simple enough to audit and fast enough for real-time feedback, making the method deployable in authentic workflows. This is not a diagnostic instrument; rather, it is a decision-support layer that foregrounds uncertainty, invites review when confidence is low, and scales across languages and settings when paired with routine calibration and fairness checks. With the provided tables, figures, and executable computation path, the approach is ready for pilot integrations and prospective validation, and it offers a practical bridge between crisp clinical rubrics and the graded, context-dependent reality of speech in the wild.

Acknowledgment

This research was partially funded by Zarqa University.

References

[1] Zadeh, L.A., 1965. Fuzzy sets. *Information and Control*. 8(3), 338–353. DOI: https://doi.org/10.1016/S0019-9958(65)90241-X

[2] Atanassov, K.T., 1986. Intuitionistic fuzzy sets. *Fuzzy Sets and Systems*. 20(1), 87–96. DOI: https://doi.org/10.1016/S0165-0114(86)80034-3

[3] Guitar, B., 2013. *Stuttering: An Integrated Approach to Its Nature and Treatment*, 4th ed. Lippincott Williams & Wilkins: Philadelphia, USA. pp. 1–520.

[4] Yairi, E.; Seery, C., 2015. *Stuttering: Foundations and Clinical Applications*, 2nd ed. Pearson: Boston, USA. pp. 1–480.

[5] Jurafsky, D.; Martin, J.H., 2009. *Speech and Language Processing*, 2nd ed. Prentice Hall: Upper Saddle River, USA. pp. 1–1024.

[6] Yogeesh, N., 2023. *Intuitionistic Fuzzy Hypergraphs and Their Operations*. Chapman and Hall/CRC: Boca Raton, USA. DOI: https://doi.org/10.1201/9781003359456

[7] Yager, R.R., 1988. On ordered weighted averaging (OWA) operators in multicriteria decision making. *Fuzzy Sets and Systems*. 18(1), 183–190. DOI: https://doi.org/10.1016/0165-0114(88)90070-9

[8] Xu, Z., 2007. Intuitionistic fuzzy aggregation operators. *IEEE Transactions on Fuzzy Systems*. 15(6), 1179–1187. DOI: https://doi.org/10.1109/TFUZZ.2006.889952

[9] Hung, W.-L.; Yang, M.-S., 2004. Similarity measures of intuitionistic fuzzy sets based on Hausdorff distance. *Fuzzy Sets and Systems*. 148(2), 215–225. DOI: https://doi.org/10.1016/j.fss.2003.10.021

[10] Grzegorzewski, P., 2004. Distances between intuitionistic fuzzy sets and some applications. *Fuzzy Sets and Systems*. 148(2), 319–328. DOI: https://doi.org/10.1016/j.fss.2003.10.009

[11] Beliakov, G.; Pradera, A.; Calvo, T., 2007. *Aggregation Functions*. Springer: Berlin, Germany. DOI: https://doi.org/10.1007/978-3-540-73721-6

[12] Yogeesh, N., 2024. Fuzzy Graph Dominance for Networked Communication Optimization. In: Sharma, V.; Balusamy, B.; Ferrari, G.; Ajmani, P. (eds.). *Wireless Communication Technologies: Roles, Responsibilities, and Impact of IoT, 6G, and Blockchain Practices*, 1st ed. CRC Press: Boca Raton, USA. pp. 30. DOI: https://doi.org/10.1201/9781003389231.

[13] Yogeesh, N., 2023. Fuzzy Clustering for Classification of Metamaterial Properties. In: Mehta, S.; Abougreen, A. (eds.). *Metamaterial Technology and Intelligent Metasurfaces for Wireless*

Communication Systems. IGI Global: Hershey, USA. pp. 200–229. DOI: https://doi.org/10.4018/978-1-6684-8287-2.ch009.

[14] Boyd, S.; Parikh, N.; Chu, E.; Peleato, B.; Eckstein, J., 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*. 3(1), 1–122. DOI: https://doi.org/10.1561/2200000016

[15] Yuan, M.; Lin, Y., 2006. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 68(1), 49–67. DOI: https://doi.org/10.1111/j.1467-9868.2005.00532.x

[16] Grabisch, M., 1996. The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research*. 89(3), 445–456. DOI: https://doi.org/10.1016/0377-2217(95)00176-X

[17] Nesterov, Y., 2004. *Introductory Lectures on Convex Optimization*. Springer: Boston, USA. DOI: https://doi.org/10.1007/978-1-4419-8853-9

[18] Niculescu-Mizil, A.; Caruana, R., 2005. Predicting good probabilities with supervised learning. *Proceedings of the 22nd International Conference on Machine Learning (ICML)*; July 2005; Bonn, Germany. pp. 625–632. DOI: https://doi.org/10.1145/1102351.1102430

[19] Ramsay, J.O., 1988. Monotone regression splines in action. *Statistical Science*. 3(4), 425–441. DOI: https://doi.org/10.1214/ss/1177012761

[20] Nijalingappa, Y.; Setty, G.D.K.; Madan, R.; Nagaraju, V.T., 2025. Directable zeros in fuzzy logics: A study of functional behaviours and applications. *AIP Conference Proceedings*. 3175(1), 020091, 10 March 2025. DOI: https://doi.org/10.1063/5.0254157

[21] Grabisch, M.; Murofushi, T.; Sugeno, M. (eds.), 2000. *Fuzzy Measures and Integrals*. Springer: Boston, USA. DOI: https://doi.org/10.1007/978-1-4615-4685-6

[22] Yogeesh, N., 2024. Solving Fuzzy Nonlinear Optimization Problems Using Evolutionary Algorithms. In: Mukherjee, G.; Basu Mallik, B.; Kar, R.; Chaudhary, A. (eds.). *Advances on Mathematical Modeling and Optimization with Its Applications*, 1st ed. CRC Press: Boca Raton, USA. pp. 20. DOI: https://doi.org/10.1201/9781003387459

[23] Dawid, A.P.; Skene, A.M., 1979. Maximum likelihood estimation of observer error-rates using the EM algorithm. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*. 28(1), 20–28. DOI: https://doi.org/10.2307/2346806

[24] Agresti, A., 2010. *Analysis of Ordinal Categorical Data*, 2nd ed. Wiley: Hoboken, USA. DOI: https://doi.org/10.1002/9780470594001

[25] Embretson, S.E.; Reise, S.P., 2000. *Item Response Theory for Psychologists*. Psychology Press: Mahwah, USA. DOI: https://doi.org/10.4324/9781410605269

[26] McGraw, K.O.; Wong, S.P., 1996. Forming inferences about some intraclass correlation coefficients. *Psychological Methods*. 1(1), 30–46. DOI: https://doi.org/10.1037/1082-989X.1.1.30

[27] Cohen, J., 1968. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*. 70(4), 213–220. DOI: https://doi.org/10.1037/h0026256

Received: July 20, 2025

[28] Dempster, A.P.; Laird, N.M.; Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*. 39(1), 1–38. DOI: https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

[29] Fleiss, J.L., 1971. Measuring nominal scale agreement among many raters. *Psychological Bulletin*. 76(5), 378–382. DOI: https://doi.org/10.1037/h0031619

[30] Gelman, A.; Hill, J., 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press: Cambridge, UK. DOI: https://doi.org/10.1017/CBO9780511790942

[31] Yogeesh, N., 2023. Fuzzy Logic Modelling of Nonlinear Metamaterials. In: Mehta, S.; Abougreen, A. (eds.). *Metamaterial Technology and Intelligent Metasurfaces for Wireless Communication Systems*. IGI Global: Hershey, USA. pp. 230–269. DOI: https://doi.org/10.4018/978-1-6684-8287-2.ch010

[32] Boucheron, S.; Lugosi, G.; Massart, P., 2013. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press: Oxford, UK. DOI: https://doi.org/10.1093/acprof:oso/9780199535255.001.0001

[33] van der Vaart, A.W., 1998. *Asymptotic Statistics*. Springer: New York, USA. DOI: https://doi.org/10.1007/978-1-4757-2778-7

[34] Rockafellar, R.T.; Wets, R.J.-B., 2009. *Variational Analysis*, 3rd ed. Springer: Berlin, Germany. DOI: https://doi.org/10.1007/978-3-642-02431-9

[35] Denneberg, D., 1994. *Non-Additive Measure and Integral*. Springer: Boston, USA. DOI: https://doi.org/10.1007/978-1-4612-4284-3

[36] Bartlett, P.L.; Mendelson, S., 2002. Rademacher and Gaussian complexities: risk bounds and structural results. *Journal of Machine Learning Research*. 3, 463–482. DOI: https://doi.org/10.1162/153244303321897690

[37] Mohammad, A. A. S., Alolayyan, M. N., Al-Daoud, K. I., Al Nammas, Y. M., Vasudevan, A., & Mohammad, S. I. (2024). Association between Social Demographic Factors and Health Literacy in Jordan. Journal of Ecohumanism, 3(7), 2351-2365.

[38] Shalev-Shwartz, S.; Ben-David, S., 2014. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press: Cambridge, UK. DOI: https://doi.org/10.1017/CBO9781107298019

[39] Mohammad, A. A., Shelash, S. I., Saber, T. I., Vasudevan, A., Darwazeh, N. R., & Almajali, R. (2025). Internal audit governance factors and their effect on the risk-based auditing adoption of commercial banks in Jordan. Data and Metadata, 4, 464.

[40] Efron, B.; Tibshirani, R.J., 1994. *An Introduction to the Bootstrap*. Chapman & Hall/CRC: New York, USA. DOI: https://doi.org/10.1201/9780429246593

[41] Mohammad, A. A. S. (2025). The impact of COVID-19 on digital marketing and marketing philosophy: evidence from Jordan. International Journal of Business Information Systems, 48(2), 267-281.

[42] Hastie, T.; Tibshirani, R.; Friedman, J., 2009. *The Elements of Statistical Learning*, 2nd ed. Springer: New York, USA. DOI: https://doi.org/10.1007/978-0-387-84858-7

Received: July 20, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[43] Mohammad, A. A. S., Mohammad, S. I. S., Al-Daoud, K. I., Al Oraini, B., Vasudevan, A., & Feng, Z. (2025). Optimizing the Value Chain for Perishable Agricultural Commodities: A Strategic Approach for Jordan. Research on World Agricultural Economy, 6(1), 465-478.

[44] Demšar, J., 2006. Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*. 7, 1–30.

[45] Mohammad, A. A. S., Mohammad, S. I. S., Al Oraini, B., Vasudevan, A., & Alshurideh, M. T. (2025). Data security in digital accounting: A logistic regression analysis of risk factors. International Journal of Innovative Research and Scientific Studies, 8(1), 2699-2709.

[46] Benjamini, Y.; Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*. 57(1), 289–300. DOI: https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

[47] Mohammad, A. A. S., Mohammad, S. I. S., Al-Daoud, K. I., Vasudevan, A., & Hunitie, M. F. A. (2025). Digital ledger technology: A factor analysis of financial data management practices in the age of blockchain in Jordan. International Journal of Innovative Research and Scientific Studies, 8(2), 2567-2577.

Received: July 20, 2025