**OpenSMILE eGeMAPS Features Training Guide**

*Understanding Your 88 Voice Analysis Features*

**🎯 Overview**

The **eGeMAPS (extended Geneva Minimalistic Acoustic Parameter Set)** extracts 88 features that represent different aspects of human voice. These features are scientifically validated for emotion recognition, personality analysis, and voice quality assessment.

**📊 Feature Categories & Meanings**

**1. F0 (Fundamental Frequency) Features - PITCH**

*10 features - Most important for voice character*

**What it measures:** The "musical note" of someone's voice - how high or low they sound

| **Feature** | **What It Means** | **Voice Archetype Clues** |
| --- | --- | --- |
| F0\_sma3nz\_amean | **Average pitch** | High = Younger/Energetic, Low = Authoritative/Mature |
| F0\_sma3nz\_stddevNorm | **Pitch variation** | High = Expressive/Dynamic, Low = Monotone/Steady |
| F0\_sma3nz\_percentile20/50/80 | **Pitch range distribution** | Shows if voice stays in narrow/wide pitch ranges |
| F0\_sma3nz\_pctlrange0-2 | **Pitch span** | Wide = Expressive speaker, Narrow = Controlled delivery |
| F0\_sma3nz\_meanRisingSlope | **Upward pitch movement** | High = Questioning/Uncertain style |
| F0\_sma3nz\_stddevRisingSlope | **Pitch rise consistency** | Variable = Natural conversation style |
| F0\_sma3nz\_meanFallingSlope | **Downward pitch movement** | High = Declarative/Confident statements |
| F0\_sma3nz\_stddevFallingSlope | **Pitch fall consistency** | Shows statement pattern regularity |

**🎭 Archetype Examples:**

* **News Anchor:** Low average F0, low variation, consistent falling slopes
* **Enthusiastic YouTuber:** Higher F0, high variation, lots of rising slopes
* **Meditation Guide:** Mid F0, very low variation, gentle falling slopes

**2. Loudness Features - ENERGY & PRESENCE**

*10 features - Key for audience engagement*

**What it measures:** How much acoustic energy and "presence" the voice has

| **Feature** | **What It Means** | **Voice Archetype Clues** |
| --- | --- | --- |
| loudness\_sma3\_amean | **Average volume/energy** | High = Commanding presence, Low = Gentle/Intimate |
| loudness\_sma3\_stddevNorm | **Energy variation** | High = Dynamic speaker, Low = Consistent delivery |
| loudness\_sma3\_percentile20/50/80 | **Energy distribution** | Shows volume consistency patterns |
| loudness\_sma3\_pctlrange0-2 | **Dynamic range** | Wide = Expressive, Narrow = Even-keeled |
| loudness\_sma3\_meanRisingSlope | **Energy build-ups** | High = Builds excitement/tension |
| loudness\_sma3\_stddevRisingSlope | **Energy rise patterns** | Shows crescendo consistency |
| loudness\_sma3\_meanFallingSlope | **Energy drops** | High = Dramatic conclusions |
| loudness\_sma3\_stddevFallingSlope | **Energy fall patterns** | Shows diminuendo style |

**🎭 Archetype Examples:**

* **Motivational Speaker:** High average loudness, wide dynamic range
* **Podcast Host:** Mid loudness, consistent energy, controlled variation
* **ASMR Creator:** Low loudness, minimal variation, gentle slopes

**3. Spectral Features - VOICE QUALITY & TIMBRE**

*18 features - What makes each voice unique*

**What it measures:** The "color" and quality of the voice - what makes it sound warm, bright, rich, etc.

**3.1 Spectral Centroid & Rolloff (4 features)**

| **Feature** | **What It Means** | **Voice Character** |
| --- | --- | --- |
| spectralCentroid\_sma3\_amean | **Voice "brightness"** | High = Bright/Sharp, Low = Dark/Warm |
| spectralCentroid\_sma3\_stddevNorm | **Brightness variation** | High = Variable tone quality |
| spectralRollOff75\_sma3\_amean | **High frequency content** | High = Crisp/Clear, Low = Muffled/Warm |
| spectralRollOff75\_sma3\_stddevNorm | **Clarity variation** | Shows tonal consistency |

**3.2 Spectral Flux (2 features)**

| **Feature** | **What It Means** | **Voice Character** |
| --- | --- | --- |
| spectralFlux\_sma3\_amean | **Voice texture changes** | High = Varied articulation |
| spectralFlux\_sma3\_stddevNorm | **Texture consistency** | Low = Smooth delivery |

**3.3 MFCC (Mel-Frequency Cepstral Coefficients) (12 features)**

*The "voice fingerprint" - most important for voice recognition*

| **Feature** | **What It Measures** |
| --- | --- |
| mfcc1-4\_sma3\_amean/stddev | **Voice timbre characteristics** |
| *These capture the unique "color" of each person's voice* |  |

**🎭 Archetype Applications:**

* **Professional Broadcaster:** High spectral centroid (bright), low flux (consistent)
* **Character Voice Actor:** High MFCC variation (versatile timbre)
* **Warm Podcast Host:** Lower centroid (warm), consistent spectral features

**4. Voice Quality Features - PROFESSIONALISM**

*6 features - Technical voice quality measures*

**What it measures:** How "clean" and professionally delivered the voice sounds

| **Feature** | **What It Means** | **Professional Quality Indicator** |
| --- | --- | --- |
| jitterLocal\_sma3nz\_amean | **Pitch steadiness** | Low = Professional control, High = Nervous/Untrained |
| shimmerLocaldB\_sma3nz\_amean | **Volume steadiness** | Low = Good breath control, High = Shaky delivery |
| HNRdBACF\_sma3nz\_amean | **Voice clarity** | High = Clean voice, Low = Breathy/Hoarse |
| logRelF0-H1-H2\_sma3nz\_amean | **Vocal cord efficiency** | Professional voicing indicator |
| logRelF0-H1-A3\_sma3nz\_amean | **Resonance quality** | Vocal technique indicator |
| F1/F2/F3amplitudeLogRelF0\_sma3nz\_amean | **Formant strength** | Articulation clarity measures |

**🎭 Professional Voice Indicators:**

* **Trained Speaker:** Low jitter/shimmer, high HNR, controlled formants
* **Casual Creator:** Higher variation, more natural irregularities
* **Nervous Speaker:** High jitter/shimmer, inconsistent quality measures

**5. Temporal Features - RHYTHM & PACING**

*14 features - How speakers use time and silence*

**What it measures:** Speaking rhythm, pauses, and timing patterns

| **Feature Group** | **What It Measures** | **Communication Style** |
| --- | --- | --- |
| **equivalentSoundLevel\_dBHL\_sma3** | **Perceived loudness** | Audience impact measure |
| **loudnessPeaksPerSec** | **Energy bursts per second** | Excitement/engagement level |
| **VoicedSegmentsPerSec** | **Speech density** | Fast = Information-dense, Slow = Contemplative |
| **MeanVoicedSegmentLength** | **Average speech chunks** | Long = Flowing, Short = Choppy |
| **StddevVoicedSegmentLength** | **Rhythm consistency** | Regular = Structured, Variable = Natural |
| **MeanUnvoicedSegmentLength** | **Average pause length** | Long pauses = Thoughtful, Short = Rapid |
| **StddevUnvoicedSegmentLength** | **Pause variation** | Consistent = Planned, Variable = Spontaneous |

**🎭 Temporal Archetype Patterns:**

* **News Anchor:** Consistent segment lengths, regular pauses, moderate speech density
* **Conversational Podcaster:** Variable segments, natural pause patterns
* **Rapid-Fire Commentary:** High voiced segments per second, short unvoiced segments
* **Meditative Speaker:** Long voiced segments, long consistent pauses

**6. Additional Quality Measures**

*28+ features - Advanced voice characteristics*

**6.1 Alpha Ratio & Hammarberg Index**

* **Alpha ratio:** Vocal effort and strain indicators
* **Hammarberg Index:** Voice roughness measures

**6.2 Slope Measures (16 features)**

* **Slope0-500Hz through Slope1500-4000Hz:** Frequency band energy distribution
* Shows voice resonance characteristics and recording quality

**6.3 Formant Features**

* **F1, F2, F3 frequency and bandwidth:** Vowel characteristics and articulation clarity

**🎯 Practical Voice Archetype Applications**

**Creator Voice Types You Can Identify:**

**1. The Authority Figure**

* **F0:** Lower average pitch, controlled variation
* **Loudness:** High average, consistent energy
* **Quality:** Low jitter/shimmer, high HNR
* **Temporal:** Measured pacing, consistent segments

**2. The Enthusiastic Host**

* **F0:** Higher pitch, high variation with rising slopes
* **Loudness:** High energy, wide dynamic range
* **Spectral:** Bright centroid, variable flux
* **Temporal:** Fast speech density, short pauses

**3. The Intimate Storyteller**

* **F0:** Mid-range, gentle falling slopes
* **Loudness:** Lower average, subtle variation
* **Spectral:** Warm (lower centroid), consistent quality
* **Temporal:** Longer segments, thoughtful pauses

**4. The Professional Educator**

* **F0:** Controlled variation, clear statement patterns
* **Loudness:** Consistent energy, purposeful dynamics
* **Quality:** Professional voice quality measures
* **Temporal:** Structured rhythm, planned pauses

**5. The Conversational Friend**

* **F0:** Natural variation, spontaneous patterns
* **Loudness:** Moderate energy, natural dynamics
* **Quality:** Slightly higher variation (more human)
* **Temporal:** Variable segments, natural pause patterns

**🔍 How to Use This for Voice Coaching**

**Identifying Voice Strengths:**

1. **Check F0 patterns** - Is the pitch range engaging?
2. **Analyze loudness dynamics** - Does energy support the message?
3. **Review quality measures** - Are there technical issues to address?
4. **Examine temporal patterns** - Is the pacing effective?

**Spotting Improvement Opportunities:**

* **High jitter/shimmer:** Breathing/tension work needed
* **Monotone F0:** Pitch variation training
* **Inconsistent loudness:** Energy control practice
* **Poor spectral quality:** Recording setup or vocal technique

**Comparing to Successful Archetypes:**

* Match feature patterns to successful creators in similar niches
* Identify which archetype fits the creator's natural style
* Recommend specific adjustments based on feature gaps

**🎼 Next Steps: Building Your Voice Archetype Library**

1. **Analyze 50+ successful creators** across different niches
2. **Cluster them by feature patterns** using your analysis tool
3. **Define 5-10 core archetypes** with specific feature profiles
4. **Create coaching recommendations** for each archetype
5. **Build a database** of archetype-to-success correlations

This will give you the foundation to offer **data-driven voice coaching** based on mathematical voice analysis rather than subjective feedback!

*This guide provides the foundation for understanding your 88 openSMILE features. Each feature represents a measurable aspect of voice that contributes to overall communication effectiveness and audience engagement.*