## **Which features were most insightful?**

## **lost\_words\_count** - incidents where speakers struggle to find specific words are strongly predictive of cognitive issues.

## **syntactic\_complexity** - With nearly equivalent correlation strength to lost words, this metric measuring sentence structure sophistication served as a powerful indicator of cognitive function.

## **speech\_rate** - The third strongest feature suggests that changes in speaking tempo significantly signal potential cognitive decline.

## **avg\_sentence\_length** - the tendency toward shorter sentences appears to be an important marker of cognitive changes.

## **word\_recall\_issues** - This related but distinct measure from lost\_words\_count captures self-corrections and word-finding difficulties, reinforcing the importance of lexical retrieval abilities in cognitive assessment.

## 

## 

## **ML Approach**

Isolation Forest was selected as the primary method for several key reasons:

1. **Unsupervised Learning**: Ideal for preliminary analysis when labeled data for cognitive decline is unavailable
2. **Anomaly Detection Focus**: Effectively identifies outliers that deviate from standard speech patterns
3. **Handling High-Dimensional Data**: Performs well with multiple speech features without requiring large datasets
4. **Interpretability**: The correlation analysis enables clear ranking of feature importance
5. **Visualization Capability**: PCA integration allowed for intuitive 2D plotting of anomalies

### **Potential Improvements:**

1. **Expanded Dataset**: Collect a larger, diverse sample set with clinical validation
2. **Longitudinal Tracking**: Implement temporal analysis to detect changes in individual speakers over time
3. **Control Groups**: Include age-matched healthy controls to establish reliable baselines
4. **Model Training:** Implement supervised learning using clinically labeled data for more accurate answers

### **Clinical Integration:**

1. **Risk Stratification System**: Develop tiered risk scoring for clinical decision support
2. **Interpretable Reports**: Create physician-friendly visualizations and summaries
3. **Integration Guidelines**: Establish protocols for use alongside traditional assessment methods

### **Other Improvements and Ideas:**

1. **Timely health checkup**: Implementation of a 60-second speech sample during routine annual physicals
2. **Testing the user**: Recent memory recall (example: "Describe what you did yesterday")
3. **Helping hand:** Connection to local cognitive health resources based on score patterns