

Departamento de Computación, FCEyN, UBA

# Procesamiento del Habla

Agustín Gravano

1er Cuatrimestre 2017

# Evaluación de sistemas TTS

- Supongamos que ya construimos un sistema de síntesis del habla. Ahora bien...
  - ¿Cómo sabemos cuán bien funciona?
  - ¿Cómo podemos compararlo contra otros sistemas?
- Enfoque tradicional: **evaluación por seres humanos**.
  - Inteligibilidad y naturalidad.

# Tests de inteligibilidad

- *Diagnose Rhyme Test (DRT)*
  - Los participantes oyen una palabra y deben decidir cuál de dos palabras fonéticamente parecidas oyeron.
  - Ejemplos: *denso* vs. *tenso*, *banda* vs. *panda*.
  - % de respuestas correctas → **puntaje de inteligibilidad**.
  - Solo para testear palabras individuales.
- *DRT Modificado (MDRT)*
  - 50 conjuntos de **6 palabras similares**.
  - Ejemplo: *basa*, *casa*, *masa*, *pasa*, *rasa*, *tasa*.
  - En **frases portadoras**: “Digo *casa* otra vez.”

# Tests de inteligibilidad

- Oraciones semánticamente **impredecibles** (SUS).
  - Template: **DET ADJ NOUN VERB DET NOUN**
  - Elegir palabras al azar en cada lugar:  
*“Los inseguros bifes cerraron los mares.”*
  - Los participantes deben **transcribir** las oraciones.
- Pruebas más cercanas a aplicaciones reales:
  - Direcciones, teléfonos, titulares de noticias, etc.
- Evaluación:
  - Porcentaje de coincidencias exactas.
  - Distancia media de edición.
    - $\text{Dist}(\text{gato}, \text{rata}) = 2$        $\text{gato} \rightarrow \text{rato} \rightarrow \text{rata}$
    - $\text{Dist}(\text{gato}, \text{perro}) = 4$        $\text{gato} \rightarrow \text{pato} \rightarrow \text{paro} \rightarrow \text{parro} \rightarrow \text{perro}$

# Naturalidad del habla sintetizada

- Puntaje medio de opinión (*Mean opinion score, MOS*)
  - Los participantes escuchan oraciones sintetizadas y evalúan la **calidad** de 1 (mala) a 5 (excelente).
  - Permite comparar varios sistemas al mismo tiempo.
- Tests “AB”
  - Los participantes escuchan varias oraciones sintetizadas con dos sistemas distintos (A y B).
  - Para cada oración, dicen cuál sistema les gusta más.
- En todos los casos, las oraciones producidas con cada sistema deben presentarse en **orden aleatorio** para evitar sesgos.

# Otras técnicas (problema abierto)

- Evaluación automática
  - Inteligibilidad: análisis espectral
  - Naturalidad?
- Evaluación basada en electroencefalografía (EEG):
  - Arndt et al. 2013, “Subjective quality ratings and physiological correlates of synthesized speech”.
- ...

