DISENTANGLING AND CONNECTING DIFFERENT PERSPECTIVES ON PROSODIC PROMINENCE
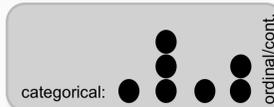
Petra Wagner[1] , Antonio Origlia[2] , Cinzia Avesani[3] , George Christodoulides[4] , Francesco Cutugno[2] , Mariapaola D'Imperio[5] , David Escudero Mancebo[6] , Barbara Gili Fivela[7] , Anne Lacheret[8] , Bogdan Ludusan[9] , Helena Moniz[10] , Ailbhe Ní Chasaide[11] , Oliver Niebuhr[12] , Lucie Rousier-Vercruyssen[13] , Anne-Catherine Simon[4] , Juraj Šimko[14] , Fabio Tesser[3] , Martti Vainio[14]

[1] Bielefeld University, [2] University of Naples Federico II, [3] CNR-ISTC, [4] Université Catholique de Louvain, [5] Aix-Marseille University, [6] Universidad de Valladolid, [7] University of Lecce, [8] Université Paris Ouest, [9] École Normale Supérieure, [10] FLUL/INESC-ID, [11] Trinity College Dublin, [12] University of Southern Denmark, [13] Université de Neuchâtel, [14] University of Helsinki
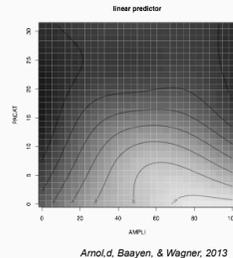
Main Perspectives

- **Functional**
  - linguistic AND paralinguistic functions
  - categorical vs. metrical vs. continuous models
  - complex, nonlinear form-function mappings
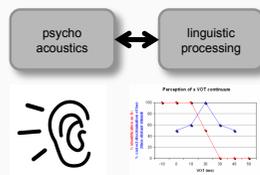 - **danger:** simplistic form-function mapping

categorical: ●●●● ordinal/cont.

- **Physical**
  - continuous rather than categorical representations
  - difficult operationalisation of prominence (functional?, psychoacoustic?)
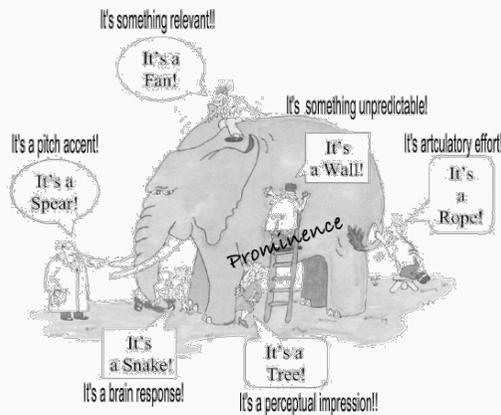 - **danger:** missing validity, e.g. due to psychoacoustic rather than communicative operationalization

linear predictor

Arnol,d, Baayen, & Wagner, 2013

- **Cognitive**
  - low-level neural pathways and psychoacoustic processing mechanisms
  - high-level linguistic processing
  - **danger**: failure to disentangle signal and function level, high- and low-level processing

psycho acoustics ↔ linguistic processing

Perception of a VOT continuum

The problem

Future Roadmap

1. Establish your perspective.
2. Distinguish bottom-up and top-down correlates.
3. Define your usage.
4. Clarify, if you follow a language-specific or universal approach.
5. Say how do your findings relate to other perspectives, based on annotations and examined domain of expression.

Connecting Perspectives

- **Less generic definitions**

  <We> say that a <linguistic entity> is prosodically prominent when it <stands out> from <its environment> (by virtue of <its prosodic characteristics>). (based on Terken (1991)

  Example: *In this study, we* say that a *word* is prosodically prominent when it *is perceived as the focus exponent* within a sen-tence by virtue of *a pitch accent.*
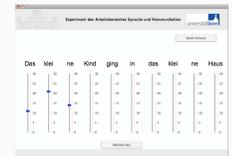
- **Cross-linguistic studies**
  - language-universality and language-specificity of signal correlates
  - L2 perception as a diagnostics for expectation-based factors

  L1 ↔ L2

- **Annotations**
  - Do continuous and categorical prominence annotations measure the same?
  - What do annotations actually measure?

- **Technical models**
  - Fully explicit
  - Agreement of human and machine annotations as diagnostic of physical and functional aspects
  - Modeling of complex function-signal interactions
  - Model evaluation by integration into technical systems, e.g. TTS

References: Arnold et al. (2013). Using generalized additive models and random forests to model prosodic prominence in German. Proc. of Interspeech 2013, 272-275. Wagner et al. (2015). Different parts of the same elephant: A roadmap to disentangle and connect different perspectives of prominence. Proc. of ICPhS 2015 | Terken, J. (1991). Fundamental frequency and perceived prominence. Journal of the Acoustical Society of America 89(4), 1768–1776.