

Prozodia niepewności

Maciej Karpiński # * & Katarzyna Klessa#

Instytut Językoznawstwa WN UAM

* ICPMiJ UAM

Prozodia w semantyce – semantyka w prozodii

Warszawa, maj 2014

Wykładniki stopnia (nie) pewności

A. Oblicza „stopnia pewności”

- wewnętrzne przekonanie mówcy co do przekazywanej treści;
- sposób realizacji wypowiedzi, który może być odczytany jako wyraz powyższego stanu;
- postrzeganie wypowiedzi przez słuchacza jako wyrażającej pewien stopień pewności.

B. Stopień pewności mogą sygnalizować wykładniki leksykalne („kwantyfikujące” wypowiedź, np. „na pewno”, „chyba”, „wydaje mi się”) – *linguistic hedges*

C. Wskazywane w literaturze prozodyczne wykładniki niepewności, m.in. (np. Smith & Clark 1993, Brennan & Williams (1995))

- Pauzy ciche i wypełnione
- Opóźnienia w powstawaniu wypowiedzi
- Wznoszące kontury intonacyjne stwierdzeń
- Wysokość głosu
- Tempo mowy



„Stopień pewności” jako parametr prozodii wypowiedzi

- Stopień pewności mówcy prawdopodobnie może być wyrażany w prozodii również **specyficznie dla danego języka**, podobnie jak obserwuje się to w przypadku prozodii emocjonalnej (zob. np. Scherer et al. 2001).
- **Kontekstualizacja** – istotny element w interpretacji znaczenia niesionego przez prozodię, w tym charakterystyka prozodyczna frazy poprzedzającej (Pon-Barry 2008; Arnold et al. 2010), jak i innych cech przekazu, również w obrębie różnych modalności i kanałów przekazu (Jun & Bishop 2013; Baumann & Hadelich 2003).



Koncepcja i metoda

- **Pytanie 1:** Czy globalne parametry prozodyczne wypowiedzi wpływają na ocenę stopnia niepewności?
- Wybór wskaźników:
 - Zakres zmian f_0
 - Średnia wysokość f_0
 - Średnie tempo wypowiedzi
- **Pytanie 2:** Czy percepcja prozodycznych wykładników niepewności pozostaje pod wpływem percepcji prozodii wypowiedzi poprzedzającej?
 - Torowanie bodźcami pozbawionymi warstwy leksykalnej



Przygotowanie materiału

- Rejestracja wypowiedzi (stwierdzeń) o podobnej długości i budowie składniowej i liczbie sylab (głos męski i żeński)
- Wybór głosu i wypowiedzi na podstawie eksperckiego odsłuchu
- Przygotowanie sygnałów (Praat, resynteza):
 - „Oryginalny” (lecz resyntetyzowany)
 - Przyspieszony (20%)
 - Spowolniony (20%)
 - Podwyższone f_0 (20%)
 - Obniżone f_0 (20%)
 - Poszerzony zakres f_0 (1,2)
 - Zredukowany zakres f_0 (0,8)
- W naturalnej mowie zmiana f_0 pociąga za sobą zmiany formantów, zaś zmiana tempa nie odbija się w równym stopniu na wszystkich segmentach. Przy w/w zakresach zmian parametrów pominięcie uwzględnienia tych zależności okazało się akceptowalne odsłuchowo.

Odsłuch 1: A vs. B

- Program AnnotationPro (Klessa et al., 2013)
- Bodźce sparowane, wymieszane pseudolosowo
 - Oryginalny + zmodyfikowany
 - Zmodyfikowany + oryginalny
 - 10 zdań x 7 wersji x 2 powtórzenia (w różnej kolejności – jak wyżej)
- Pytanie: W którym przypadku wypowiedź brzmiała pewniej?
- Odsłuchujący: 16 osób (studenci WN), ostatecznie uwzględniono dane od 14 odsłuchujących

AnnotationPro – program dostępny bezpłatnie do celów naukowych, umożliwia transkrypcję, anotację i jej analizę oraz prowadzenie eksperymentów odsłuchowych: annotationpro.org.



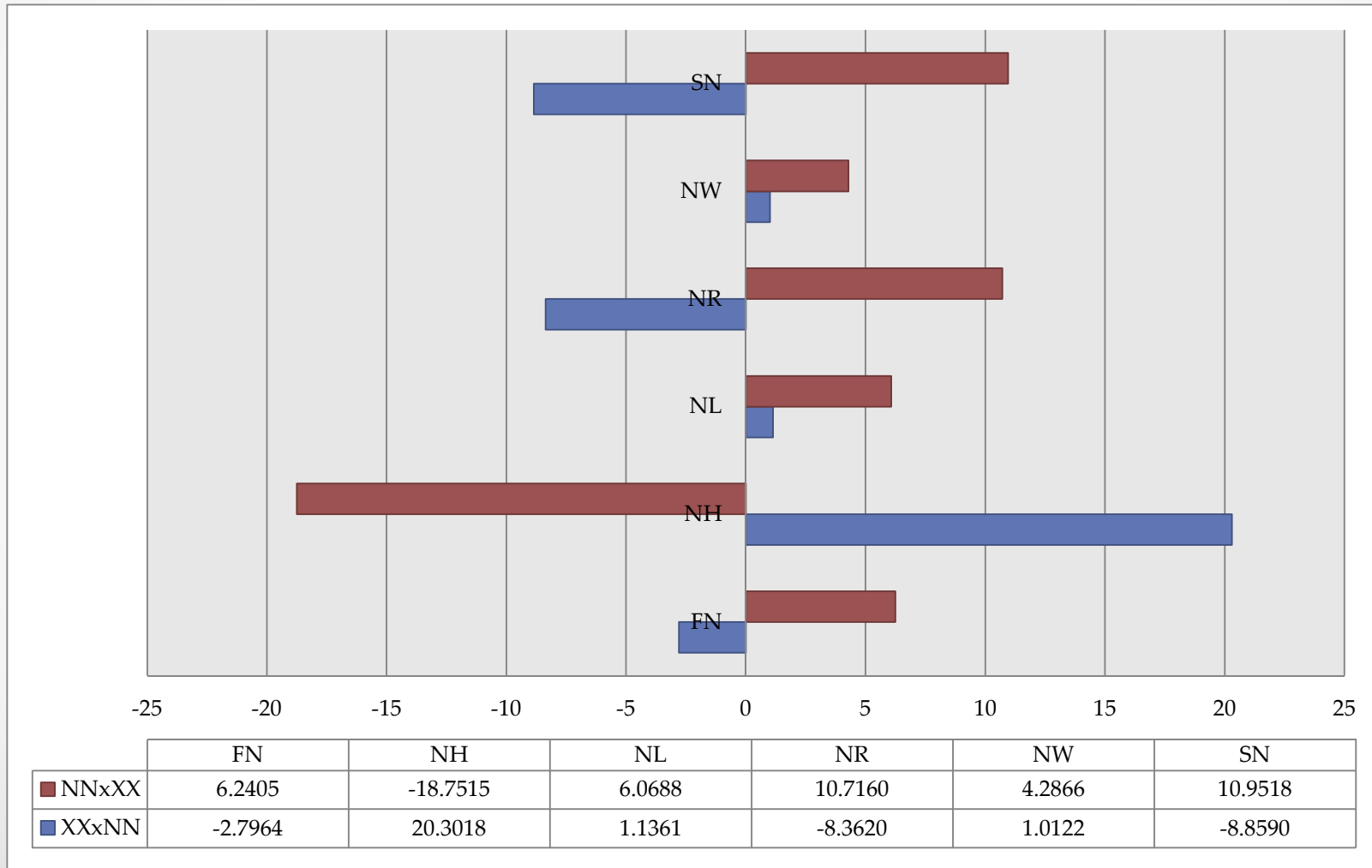
Przebieg

The screenshot displays the Annotation Pro 1.9.1.0 software interface. The main window is titled "A0001_N_N_A0001_F_N.wav | Annotation Pro 1.9.1.0". The interface is divided into several panels:

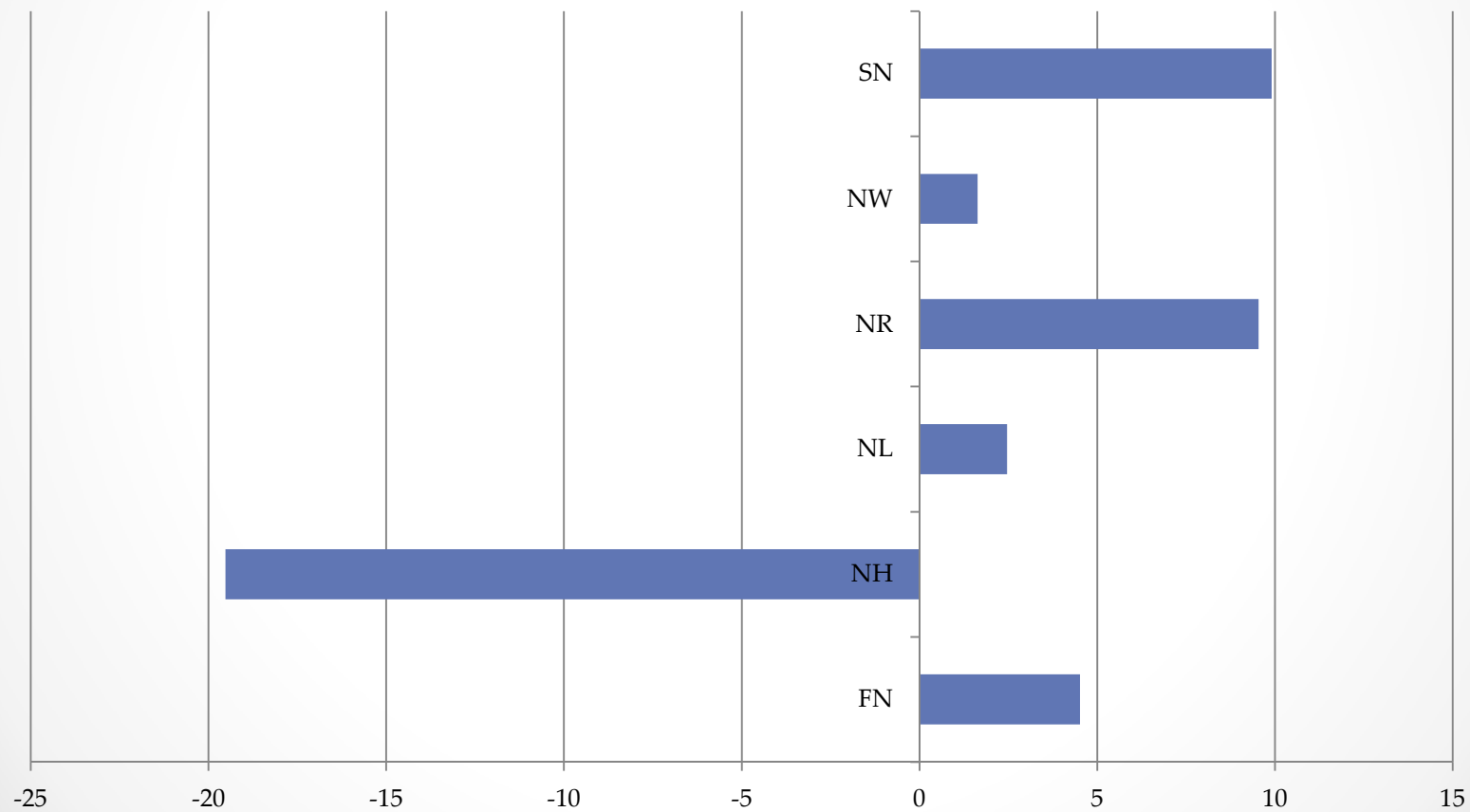
- Workspace:** A list of audio files and annotations. The "Ungruiped" section contains files like "A0001_F_N_A0001_N_N...", "A0001_N_H_A0001_N...", "A0001_N_L_A0001_N_N...", "A0001_N_N_A0001_F_N..." (highlighted), "A0001_N_N_A0001_N...", "A0001_N_N_A0001_N_L...", "A0001_N_N_A0001_N...", "A0001_N_N_A0001_N...", "A0001_N_N_A0001_S...", "A0001_N_R_A0001_N...", "A0001_N_W_A0001_N...", "A0001_S_N_A0001_N...", "A0003_F_N_A0003_N_N...", "A0003_N_H_A0003_N...", "A0003_N_L_A0003_N_N...", "A0003_N_N_A0003_F_N...", "A0003_N_N_A0003_N...", "A0003_N_N_A0003_N_L...", "A0003_N_N_A0003_N...", "A0003_N_N_A0003_S...", "A0003_N_R_A0003_N...", "A0003_N_W_A0003_N...", "A0003_S_N_A0003_N...", "A0004_F_N_A0004_N_N...", "A0004_N_H_A0004_N...", "A0004_N_L_A0004_N_N...", "A0004_N_N_A0004_F_N...", "A0004_N_N_A0004_N...", "A0004_N_N_A0004_N_L...", "A0004_N_N_A0004_N...", "A0004_N_N_A0004_N...", "A0004_N_N_A0004_S...".
- Audio:** A waveform view showing two audio channels. The time axis ranges from 00:00:00 to 00:00:04:551. Below the waveform is an "Annotation" layer with a green segment labeled "23;11".
- Feature:** A diagram showing a timeline with markers labeled "A" and "B". Below the timeline is a sequence of numbers: 2, 1, 0, 1, 2.
- Properties:** A panel for editing the selected annotation. It includes fields for "Label" (set to "-19;9"), "Speaker", "Name" (with "Par. 1" and "Par. 2" options), and "Marker". There are also "Fore Color" and "Back Color" buttons.

At the bottom of the window, the status bar displays: "Audio: 16 bit PCM: 44kHz 1 channels | Annotation: 44 kHz | Rate N: 8.82 seg/s | Rate G: 8.82 seg/s | Klessa".

Odstuch 1: Wyniki



Odstuch 1: Wyniki



Odśłuch 2: Torowanie

Procedura torowania:

Bodziec torujący (pryma) **A** poprzedza bodziec torowany (główny) **B**. Badana jest zmiana reakcji uczestników badania na ten bodziec w zależności od braku/obecności/różnego charakteru prymy.

Zadanie: Określ stopień pewności wypowiedzi odtworzonej po sygnale (wcześniej przedstawiono przykładowy „sygnał”, stanowiący prymę)

6 zdań x 5 wersji x 5 prym = 150 bodźców

Odśłuch 2: Torowanie

- Bodziec torujący (pryma) - pięć typów, opartych na tym samym sygnale oryginalnym
 - N - Sygnał oryginalny (resynteza)
 - H - Podwyższone f_0
 - L - Obniżone f_0
 - W - Poszerzony zakres f_0
 - R - Zawężony zakres f_0

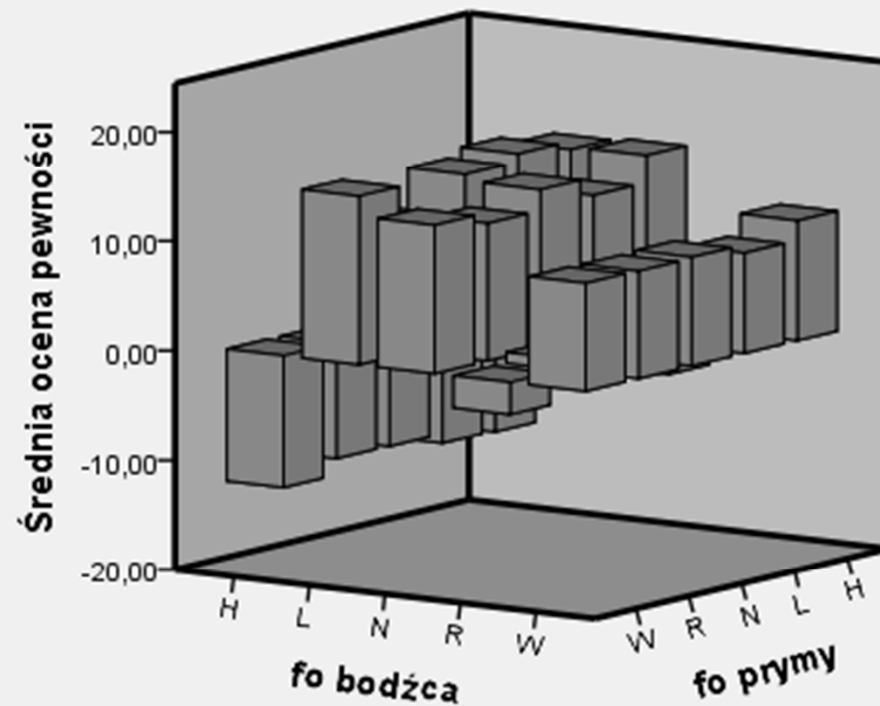
 - **Resynteza tonu podstawowego w postaci „hum” (Praat) – pryma pozbawiona zawartości leksykalno-składniowej**
 - Rezygnacja z bodźców ze zmienionym tempem ze względu na brak struktury segmentalnej w bodźcu

Odsłuch 2: Przebieg

- $\{N_N|N_L|N_H|N_R|N_W\} + \{1 \text{ sek. Ciszy}\} + \{N_N|N_L|N_H|N_R|N_W\}$

The screenshot displays the Annotation Pro 1.9.1.0 software interface. The main window shows a waveform of an audio file with a selected segment from 00:00:01,603 to 00:00:02,404. Below the waveform, an annotation layer is visible with a segment labeled '-37;11'. The interface includes a menu bar (File, Edit, View, Statistics, Tools, Plugins, Help), a toolbar with various playback and editing controls, and a workspace on the left listing audio files. On the right, there is a 'Feature' panel with a timeline and labels 'wypowiedź niepewna' and 'wypowiedź pewna', and a 'Properties' panel with fields for Label, Speaker, Fore Color, Back Color, Marker, Name, and Paragraphs (Par. 1, Par. 2). The status bar at the bottom indicates 'Audio: 16 bit PCM: 44kHz 1 channels', 'Annotation: 44 kHz', 'Rate N: 0.94 seg/s', and 'Rate G: 0.94 seg/s'.

Odśłuch 2: Wyniki



Podsumowanie wyników

- Wyniki odsłuchu 1
 - Wypowiedź realizowana wyższym głosem jest postrzegana jako „mniej pewna”
 - Wypowiedź wolniejsza (SN) oraz bardziej płaska intonacyjnie (NR) postrzegana jako „pewniejsza”
- Wyniki odsłuchu 2
 - Wpływ prymy na ocenę bodźców jest znikomy/nieistotny
 - Dla oceny stopnia pewności istotna jest kategoria samego bodźca
 - Wypowiedź realizowana wyższym głosem jest postrzegana jako „mniej pewna” (zgodnie z wynikami odsłuchu 1)
 - Wypowiedź płaska intonacyjnie postrzegana jako mniej pewna (inaczej niż w przypadku odsłuchu 1 – efekt wymaga wyjaśnienia)
 - Wypowiedzi o obniżonej częstotliwości podstawowej oraz „standardowe” (niemodyfikowane) były percypowane jako charakteryzujące się największym stopniem pewności

Dyskusja i dalsze badania

- Problemy
 - „zagęszczenie bodźców” – wobec możliwości percepcyjnych uczestników
 - laboratoryjny charakter badania
 - charakter manipulacji bodźcami (bez uwzględnienia niektórych zjawisk związanych ze zmianą tempa mowy i wysokości głosu)
- Planowane modyfikacje
 - Powiększenie grupy badanych
 - Inna procedura odsłuchowa (mniejsze grupy sygnałów, dłuższe przerwy)
 - Subtelniejsze manipulacje parametrami sygnałów (np. krokowe zmiany wartości)



Dziękuję za uwagę



Literatura (wybrane pozycje)

- **Borràs-Comes, J., Roseano, P., del Mar Vanrell, M., Chen, A., Prieto, P. 2011.** Perceiving uncertainty: facial gestures , intonation, and lexical choice. Gespin 2011 Proceedings, Bielefeld.
- **Jun, S.-A. & Bishop, J. 2013.** Implicit Phrasing and Prominence Influence Relative Clause Attachment: Evidence from Prosodic Priming, Talk presented at the WCCFL 31, Arizona State University.
- **Karpiński, M. 2012.** The Boundaries of Language: Dealing with Paralinguistic Features. *Lingua Posnaniensis*, vol. LIV (2)/2012. PL ISSN 0079-4740, ISBN 978-83-7654-252-2, pp. 37-54.
- **Klessa, K., Karpiński, M., Wagner, A., 2013.** Annotation Pro - a new software tool for annotation of linguistic and paralinguistic features. Proc. of the Tools and Resources for the Analysis of Speech Prosody (TRASP) Workshop, Aix en Provence, 2013 (do pobrania: www.annotationpro.org).
- **Klessa, K., Wagner, A., Oleśkiewicz-Popiel, M., Karpiński, M. 2013.** "Paralingua" - a new speech corpus for the studies of paralinguistic features. *Procedia - Social and Behavioral Science*. Vol. 95. pp. 48-58. Elsevier ISSN: 1877-0428.
- **Liscombe, J., Hirschberg, J. and Venditti, J., 2005.** Detecting certainness in spoken tutorial dialogues, Proceedings of Eurospeech 2005, Lisbon, Portugal.
- **Ohala, J. 1984.** An ethological perspective on common cross - language utilization of F0 of voice. *Phonetica*, 41.1 - 16.
- **Pickering, M. J., & Ferreira, V. S. 2008.** Structural priming: A critical review. *Psychological Bulletin*, 134, 427-459.
- **Pon-Barry, H. 2008.** Prosodic Manifestations of Confidence and Uncertainty in Spoken Language. Proceedings of InterSpeech 2008.
- **Scherer, K. R., Banse, R., and Wallbott, H. G. 2001.** Emotion inferences from vocal expressions across languages and cultures. *Journal of Cross-Cultural Psychology* 32 (1), 76-92.
- **Wesson, C. & Pulford, B. D. 2009.** Verbal expressions of confidence and doubt. *Psychological Reports*, 105, 151-160.