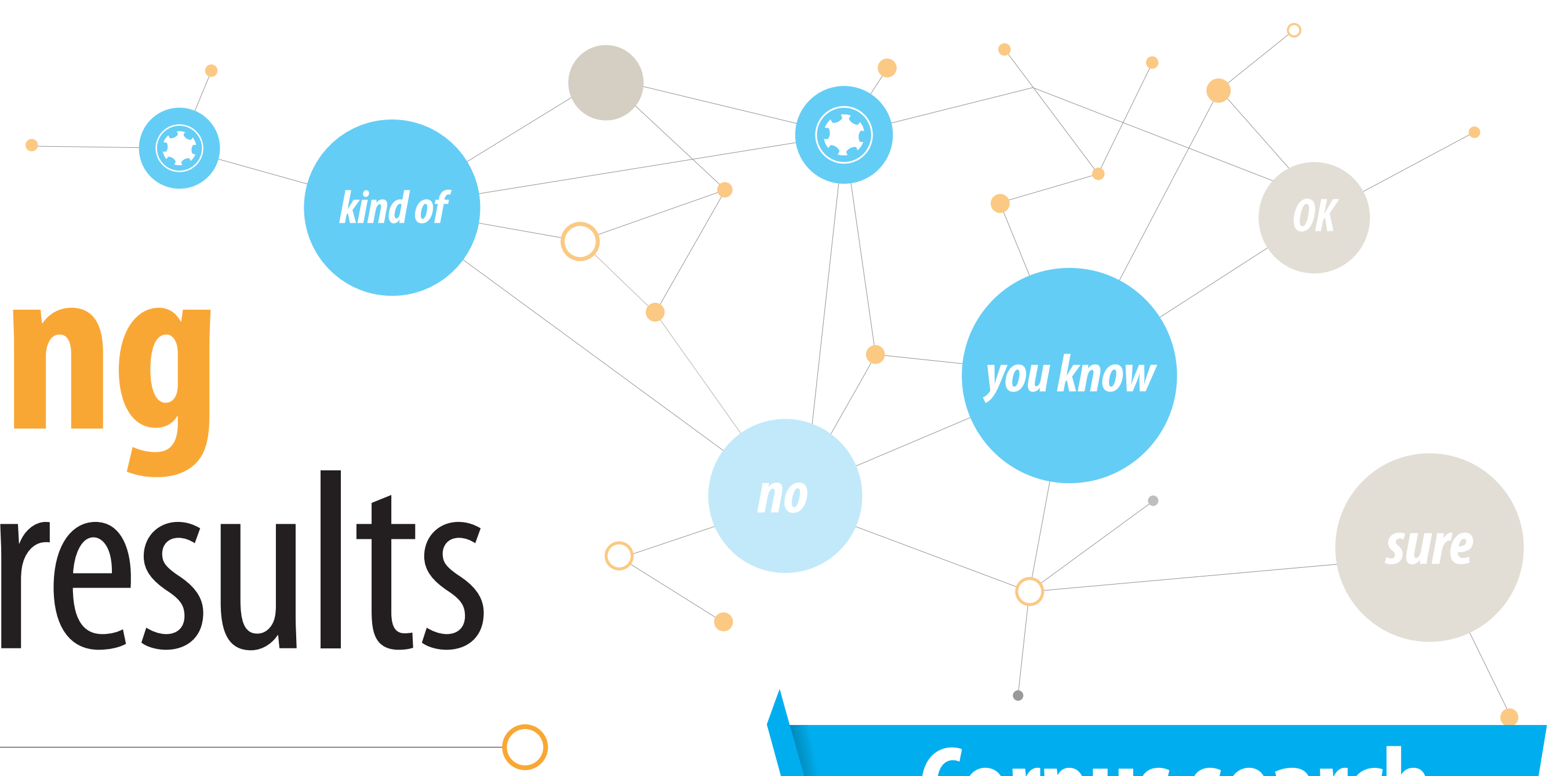


Prosody-based clustering of spoken corpus search results



Piotr Pezik, University of Łódź, CLARIN-PL
piotr.pezik@uni.lodz.pl

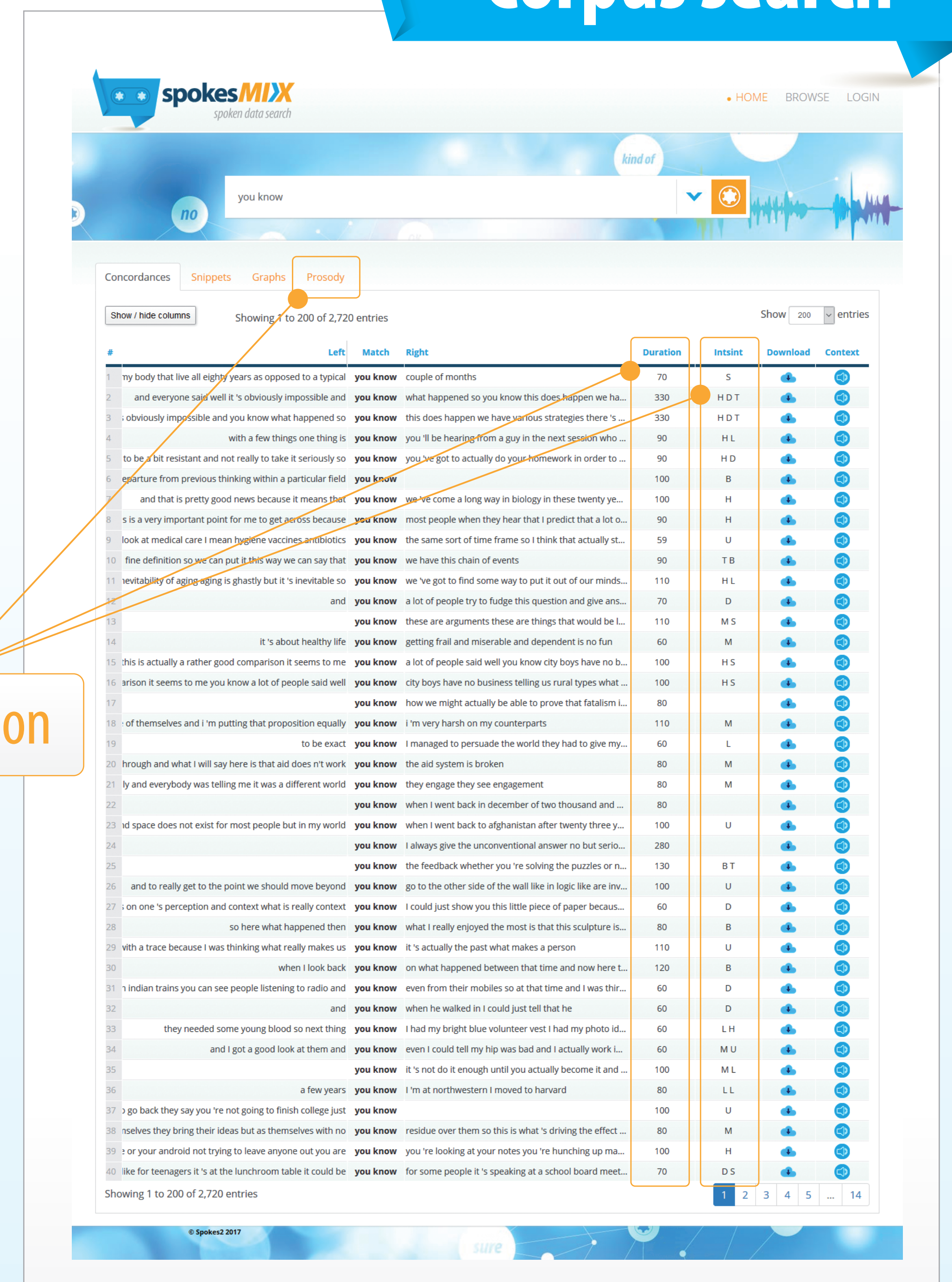
Corpus search

Rationale

The role of prosodic features of speech in construction of meaning and discourse structure is an active area of research (Wichmann et al. 2009). At the same time, there are relatively few tools for studying prosodic phenomena in reference corpora of conversational and read speech.

The technique presented in this poster facilitates analyses of corpus search results according to selected features of the speech signal computed for word spans matching user queries. For example, given a query for the phrase *you know*, it is possible to automatically label pitch patterns of its instances and thus distinguish between rising vs. falling pitch contours or relatively high- vs. relatively low-pitched instances of *you know* in the corpus. The assumption behind this approach is that even a simple sorting of search results by such labels may lead to the discovery of prosody-marked functions and meanings of words and multiword expressions.

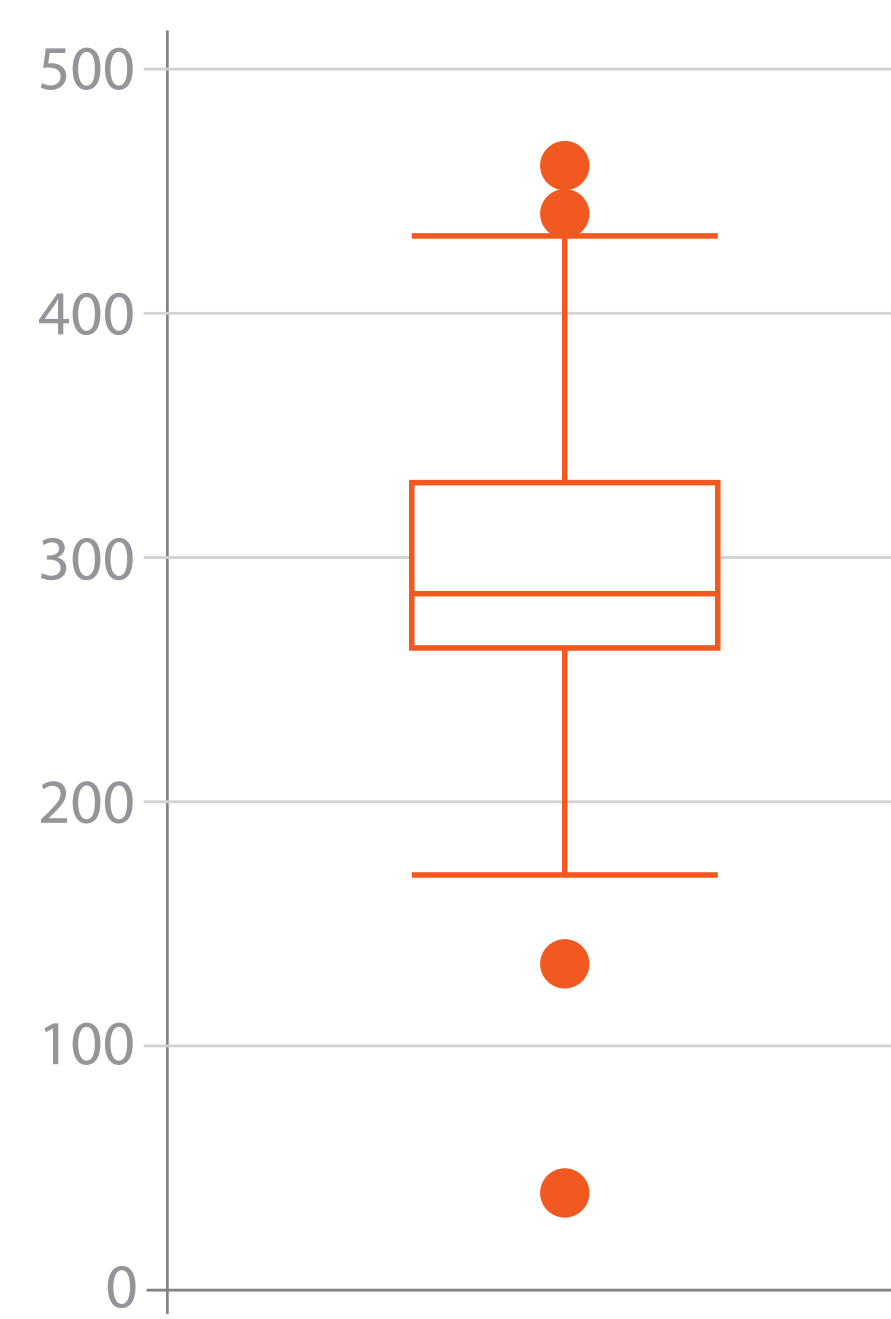
Aggregation



Span duration

Statistics Value

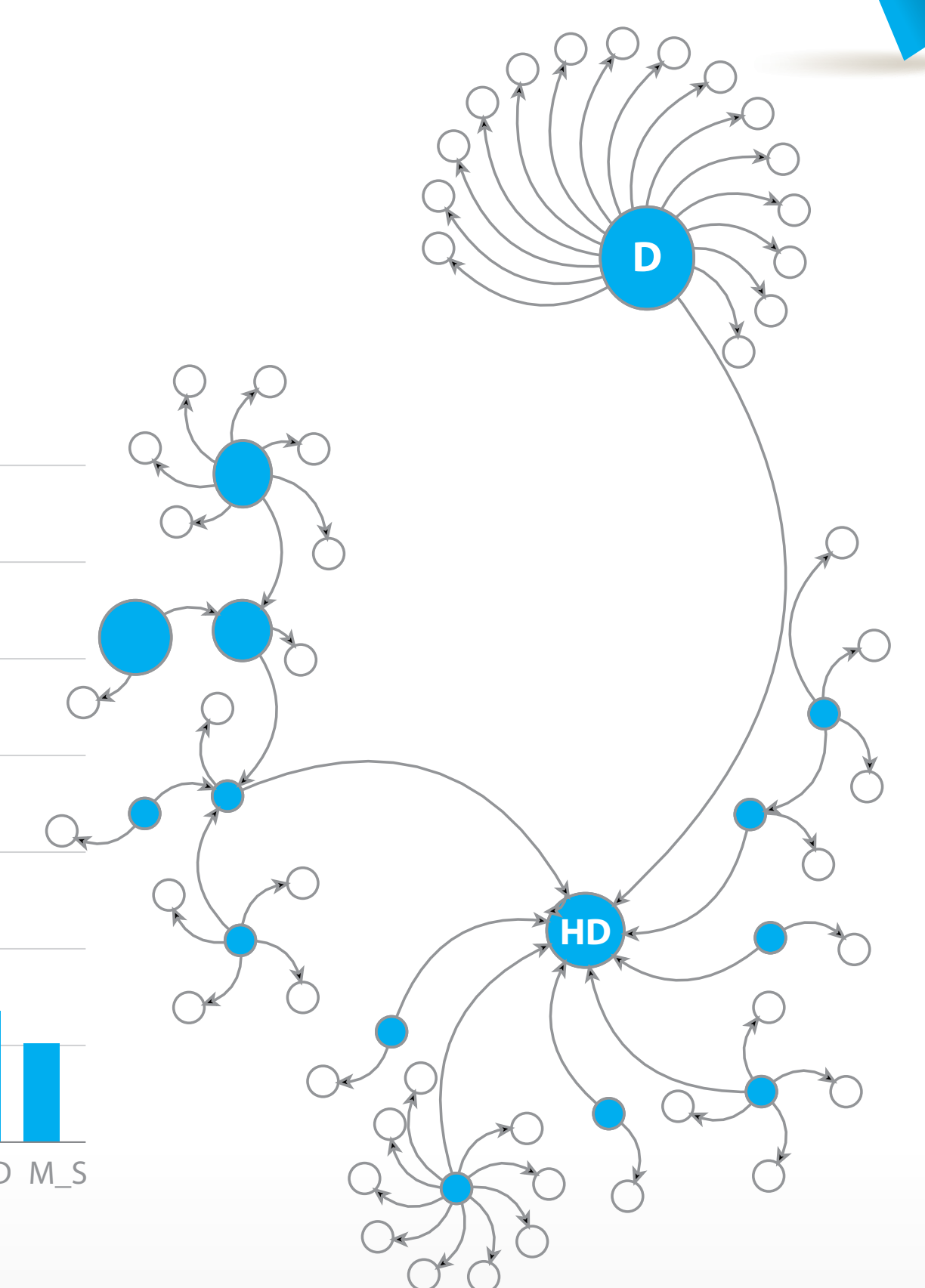
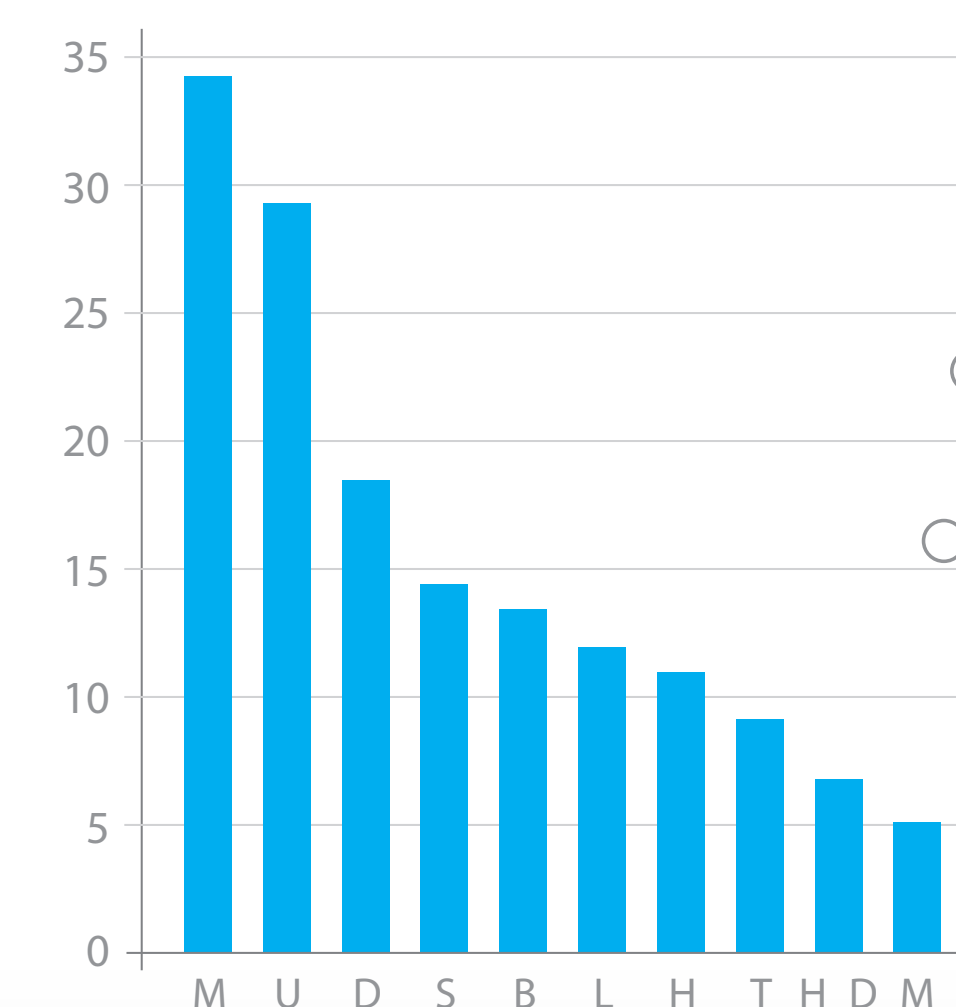
Sample size	225
Missing values	0
Minimum	59
1st Quartile	80
Median	90
3rd Quartile	110
Max	430
Mean	100.52
Variance	2401.16
St. dev	49.00



Intonation codes

Intsint codes Count

1	M	34
2	U	29
3	D	18
4	S	14
5	B	13
6	L	12
7	H	11
8	T	9
9	H_D	7
10	M_S	5



Qualitative analysis



How it works

1. Corpus search results in Spokes (cf. Pezik 2015) are aggregated on duration and intonation values measured for individual spans.
2. Prosodic contours of matched spans are encoded according to the Momel and Intsint schemes (Hirst & Espesser 1993):
 - Global codes: **Top**, **Middle**, **Bottom**
 - Relative codes: **Higher**, **Upstepped**, **Same**, **Downstepped**, **Lower**
3. In addition to simple rankings, it is possible to cluster matched spans by their intonation patterns, possibly also combining them with duration values as a separate dimension.

References

- Hirst, D. and Espesser, R. 1993. Automatic modeling of fundamental frequency using a quadratic spline function. Travaux de l'Institut de phonétique d'Aix-en-Provence, 15, 71-85.
- Pezik, P. 2015 "Spokes – a Search and Exploration Service for Conversational Corpus Data." In Selected Papers from the CLARIN 2014 Conference, October 24-25, 2014, Soesterberg.
- Wichmann, A., Dehé, N. and Barth-Weingarten D. 2009. Where Prosody Meets Pragmatics: Research at the Interface. In Where Prosody Meets Pragmatics, 1–20. Brill.

Demo

1. Go to <http://pelcra.clarin-pl.eu/spokes2-web>.
2. Type a query and restrict results to "Tedlium".
3. Click on the "Prosody" tab.