Probing visualizations of neural word embeddings for lexicographic use

Ágoston TÓTH, Esra ABDELZAHER

University of Debrecen, Faculty of Humanities, Department of English Linguistics

E-mail: toth.agoston@arts.unideb.hu, esra.abdelzaher@gmail.com

UNIVERSITY of DEBRECEN

T·U·D·A
UNIVERSITY OF DEBRECEN
FACULTY OF HUMANITIES
SCHOLARLY FUND

Agenda, aims

- to visualize the distributional features of occurrences of selected headwords in dictionary examples to see if lexicographical sense delineation is reflected in the distributional data

- to check if visualizations of BERT data are useful for assisting manual sense delineation

- to better understand the distributional information stored in BERT representations

Motivation | the power of Distributional Semantics

Meaning is a function of distribution (Harris, 1954)

In practice,

- the more similar the context, the more similar the meaning
- morphological, semantic, etc. paradigms can be reproduced using vector arithmetics

Phases of Distributional Semantics:

1. Non-ANN phase (“count methods”)

2. *2013*- ANN-based static word embeddings

e.g. word2vec

3. *2018*- contextualized ANN-based token embeddings:

non-generative: ELMo, BERT, etc.

generative: GPT, T5, etc.

Methods | Data Collection

Example sentences were collected for 4 words: *full, mouth, risk* and *sound*:

- all matching examples from the online *Oxford Learner's Dictionaries* (http://www.oxfordlearnersdictionaries.com)

- 1000 randomly selected corpus sentences for each word from the *British National Corpus* via http://www.sketchengine.eu

Methods | BERT embeddings

- We produced BERT embeddings for the headword in each example sentence by running the neural network; the neural activations for the target words were extracted and saved for visualization.
- Language Model: the largest pretrained BERT LM from Huggingface, *bert-large-uncased* (https://huggingface.co/bert-large-uncased)
- LM size: 336 million pre-trained parameters with 24 layers and 16 attention heads
- word embedding size: 1024 floating point numbers per embedding

Methods | BERT embeddings

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	
1	Sec	Sentence	Unit#0	Unit#1	Unit#2	Unit#3	Unit#4	Unit#5	Unit#6	Unit#7	Unit#8	Unit#9	Unit#10	Unit#11	Unit#12	Unit#13	Unit#14	Unit#15	Unit#16	Unit#17	Unit#18	Unit#19	Unit#20	Unit#21	Unit#22	Uni
2	1	×OD1×She opened her mouth to say something.	-1.92338	-0.57999	-0.81672	0.013772	-0.25877	-0.10113	-0.34971	0.021202	0.563403	1.232823	-0.42498	-0.21426	-0.84699	-0.2062	0.268531	0.365866	0.160868	-0.87528	0.486357	-0.71282	-0.0578	-0.42166	-1.10357	-0
3	1	×OD1×His mouth twisted into a wry smile.	-0.84353	-0.20426	-0.86526	-0.46331	-0.17053	0.213446	-0.58634	-0.05767	0.190861	1.25925	-0.19409	-0.35384	-0.44052	-0.61443	0.924573	0.271231	-0.34372	0.15834	0.530733	-0.91381	-0.31297	-0.08441	-0.06656	0.4
4	1	×OD1×Don't talk with your mouth full.	-1.47395	-0.41508	-0.44285	-0.24847	-0.1912	0.225378	-0.01788	0.570562	0.372587	0.839025	-0.16348	-0.20358	-0.78552	-0.5304	0.932828	-0.18406	-0.1819	-0.23539	0.561009	-1.01298	0.185537	-0.28952	-0.95022	-0
5	1	×OD1×The creature was foaming/frothing at the mouth.	-0.58245	-0.28124	-0.12902	-0.494	-0.16358	-0.03265	0.084395	0.316399	-0.22958	0.998992	-0.6834	-0.01081	-0.63671	-0.61716	0.497619	0.220153	0.112991	-0.38152	0.434511	-0.90718	0.146692	-0.25428	-0.60248	-0
6	1	×OD1×Cover your mouth when you cough.	-1.69553	-0.09299	0.158295	0.189177	-0.78173	0.041425	-0.0044	0.163477	-0.09538	1.151161	-0.31014	-0.04225	-0.66941	-0.55502	0.168129	0.612486	0.179191	-0.15891	0.118924	-0.69534	0.178991	0.125525	-0.95487	-0
7	1	×OD1×Rinse your mouth thoroughly with water.	-1.21488	-0.23636	-0.04907	-0.07036	-0.55905	-0.00195	0.465448	0.333199	0.069441	1.059498	0.356908	-0.38242	-0.67213	-0.72558	0.500171	0.5422	0.345466	-0.3552	-0.01416	-0.81019	0.463702	-0.19286	-0.88029	-0
8	1	×OD1×A cool smile played across her mouth.	-0.81855	-0.09352	-0.54884	-0.43545	-0.28202	0.016202	-0.08375	-0.17314	-0.23316	1.002504	0.013239	-0.34665	-0.24712	-0.58104	0.175159	0.197156	-0.63032	0.062073	0.346941	-0.75001	-0.32249	-0.15742	-0.01739	-0
9	1	×OD1×A smile played around his strong mouth.	-0.57073	-0.12001	-0.83702	-0.49859	-0.178	0.011814	-0.27115	-0.14874	0.234403	1.262916	-0.54415	-0.3905	-0.19458	-0.31615	0.597124	0.126927	-0.45203	0.323202	0.290225	-0.85929	-0.39679	0.00091	0.052696	0.0
10	1	×OD1×A tight mouth was the only sign of her nerves.	-0.85344	0.124154	-0.50282	-0.42683	-0.47023	0.205279	-0.37605	0.154819	0.002069	1.426979	-0.42187	-0.69614	-0.66132	-0.47364	0.201446	0.219378	-0.27094	-0.48661	0.423517	-0.90115	0.287359	-0.30739	-0.58226	0.0
11	1	×OD1×Good mouth care is very important when you are h	-0.76129	-0.05973	-0.26132	-0.09327	-0.40448	0.134686	0.223892	0.021125	-0.18381	0.852877	-0.2365	-0.38648	-0.25089	-0.38304	0.344595	0.137519	0.259275	-0.31575	0.010419	-0.48608	0.027095	-0.46221	-0.64036	0.0
12	1	×OD1×He began to stuff his mouth with pasta.	-1.5603	-0.23903	-0.42521	-0.18165	-0.58804	0.277995	0.134776	0.55495	0.421135	0.937742	-0.08346	-0.40599	-0.52008	-0.64912	0.440315	0.440134	-0.12988	-0.38282	-0.02826	-0.84354	0.419516	-0.06632	-1.35812	-0
13	1	×OD1×He coughed as the blood filled his mouth.	-0.6701	0.068409	0.108762	-0.28542	-0.57624	-0.08484	0.299774	-0.05794	0.231223	1.185592	-0.06194	-0.19988	-0.71253	-0.25292	0.303237	0.409187	0.139207	-0.36129	0.250282	-0.73842	0.303269	-0.39572	-0.9492	-
14	1	×OD1×He covered his mouth to hide his yawn.	-1.43886	-0.08641	-0.04479	0.061276	-0.87365	0.234403	-0.35959	-0.00479	0.462444	1.564879	-0.4446	0.092227	-0.92354	-0.33919	0.45486	0.393009	0.09848	-0.34889	0.340137	-0.77424	0.088681	-0.20598	-0.8803	0.0
15	1	×OD1×He wiped his greasy mouth on his sleeve.	-0.83464	-0.49037	-0.08088	0.019463	-0.46088	0.333183	0.238764	0.341348	0.193705	1.387084	-0.25896	-0.08732	-0.91089	-0.89915	0.431564	0.553958	0.580302	-0.54968	0.373366	-0.87918	0.113394	-0.31398	-0.65005	0.0
16	1	×OD1×Her mouth curved into a smile.	-0.80613	-0.06022	-0.95013	-0.44658	-0.20902	0.225637	-0.29163	0.185723	0.072798	0.765397	-0.30248	-0.34337	-0.15791	-0.25762	0.884377	0.107261	-0.2861	0.189481	0.389909	-0.82257	-0.36522	0.016791	-0.35438	0.1
17	1	×OD1×Her mouth suddenly set in a determined line.	-0.87063	-0.27557	-0.37738	-0.41786	-0.18526	0.327939	-0.54443	0.021382	-0.35155	0.984733	0.071187	-0.37228	-0.27648	-0.57025	0.503646	0.162327	-0.59562	-0.39692	0.253896	-1.001	0.073254	-0.12686	0.061772	0.2
18	1	×OD1×His mouth compressed into a thin, hard line.	-0.45282	-0.19998	-0.44098	-0.43712	-0.41154	0.449575	-0.42495	0.047547	-0.13686	1.363066	-0.27189	-0.3759	-0.40268	-0.92139	0.639405	0.200195	-0.35858	-0.35634	0.292817	-1.0592	0.034575	-0.30001	0.089983	
19	1	×OD1×His mouth lifted in a wry smile.	-1.15634	-0.33669	-0.80876	-0.63072	0.041226	0.394436	-0.68275	0.076647	0.018364	1.378343	-0.28563	-0.27257	-0.15581	-0.60564	1.067816	0.407116	-0.2364	0.3075	0.741073	-0.92334	-0.20709	-0.09791	-0.14026	0.5
20	1	×OD1×His mouth widened to a smile.	-0.8233	-0.00981	-0.90419	-0.6311	-0.11871	0.212881	-0.36976	0.136867	0.119987	0.744976	-0.4057	-0.20056	-0.14287	-0.11356	0.844496	-0.00899	-0.3991	0.051094	0.472921	-0.54499	-0.07517	-0.00962	-0.31651	0.1
21	1	×OD1×I could taste blood in my mouth.	-0.95154	-0.0894	-0.36702	-0.57394	-0.73049	-0.01435	0.626404	0.068963	0.180172	0.881084	0.267702	-0.186	-0.70513	-0.7715	0.260782	0.673633	-0.34675	-0.46352	0.778223	-0.716	-0.14204	-0.68055	-1.08341	-1
22	1	×OD1×I was so thirsty my tongue was sticking to the roof c	-1.5136	-0.02881	-0.16866	-0.35852	-0.90987	0.183986	0.384561	0.139789	0.401905	1.252623	0.15805	-0.32031	-0.99107	-0.92888	0.485538	0.135037	0.149744	0.33946	1.103267	-0.70004	0.150278	-0.66621	-0.87621	-1
23	1	×OD1×My mouth started watering when I smelled the foo	-0.99949	-0.08359	-0.87463	-0.52943	-0.22063	0.168161	0.151061	0.710101	0.349825	0.32195	0.029899	-0.36609	-0.6019	-0.40135	0.350096	-0.26432	-0.32599	-0.0629	0.379094	-0.92618	0.148157	-0.17279	-0.96165	-
24	1	×OD1×She put her hand over her mouth to stifle the cougl	-1.40193	0.25207	-0.1323	0.14765	-0.79616	0.114024	0.168806	0.116237	0.011204	1.29204	-0.19007	-0.01759	-0.91058	-0.2815	0.380685	0.265324	0.132838	-0.03272	0.32984	-0.53855	0.435339	-0.22271	-1.06734	-
25	1	×OD1×Suddenly a hand cupped her mouth.	-1.28596	-0.08496	-0.51653	-0.1126	-0.49682	-0.00094	0.085313	0.163529	-0.40226	0.587064	-0.11137	-0.15149	-0.58434	-0.13094	0.00459	0.129679	0.637826	0.280603	0.028446	-0.65013	0.318993	-0.50495	-0.79338	-0
26	1	×OD1×The corners of her mouth turned up in a slight smil	-0.93211	-0.24092	-0.71767	-0.42456	0.15828	0.227332	-0.51477	-0.02649	0.08163	1.085336	-0.41445	-0.53015	-0.2989	-0.29507	0.892917	0.0202	-0.38419	0.248495	0.549564	-1.02544	-0.35192	0.031535	-0.33174	0.2
27	1	×OD1×The hot coffee burned her mouth.	-0.96835	0.017189	0.012659	-0.35617	-0.49413	-0.0562	0.361481	0.415918	0.157395	1.101655	0.097159	-0.31295	-0.83437	-0.7565	0.419754	0.30211	0.119938	-0.17457	0.64515	-0.92551	0.077107	-0.30692	-1.12271	-0
28	1	×OD1×There was blood trickling from the corner of his mc	-0.95469	-0.55333	-0.30734	-0.52481	-0.68502	-0.07053	0.153381	0.053977	0.068224	1.695402	-0.47046	-0.47589	-0.7673	-0.49689	0.40804	0.587847	0.334492	-0.40845	0.60798	-0.68212	-0.1655	-0.15186	-0.73627	-0
29	1	×OD1×There were lines of tension about his mouth.	-0.69541	-0.17347	-0.42984	-0.59849	-0.44829	-0.1697	-0.54049	0.054927	-0.16675	1.567305	-0.29453	-0.62222	-0.57886	-0.68545	0.750656	0.374357	-0.42825	-0.38222	0.576791	-0.93575	0.078811	-0.26828	-0.05366	-0
30	2	×OD2×Now there would be another mouth to feed.	-0.90851	-0.15787	-0.45865	-0.21655	-0.29434	0.046581	0.36372	0.363131	-0.07377	0.139783	-0.04588	-0.01797	-0.11296	-0.45845	0.194421	0.44486	0.258809	0.304734	-0.12614	-0.79785	-0.31898	-0.48501	-0.60611	-0
31	3	×OD3×A number of industries sprang up around the mout	-0.81719	0.0392	-0.47801	-0.21128	-0.13653	-0.12704	0.44828	0.240459	-0.14705	0.95911	-0.72292	-0.37019	-0.63851	-0.18723	0.60696	0.171166	-0.17801	-0.14466	0.233757	-0.08809	-0.02524	-0.49124	-0.41341	-
32	4	×OD4×Up ahead was the tunnel mouth.	-0.63092	-0.03469	-0.45633	0.320619	-0.50811	0.19754	0.093574	0.343336	-0.2274	0.683108	-0.13603	-0.23261	-0.3396	-0.49935	0.175403	0.372289	0.24504	0.155394	0.353914	-0.5901	0.159117	-0.38166	-0.67982	-0
33	4	×OD4×They drew nearer to the mouth of the cave.	-0.62985	-0.06855	-0.44887	0.198702	0.008689	0.166648	0.226382	0.552116	-0.31677	0.820151	-0.21385	-0.23275	-0.28873	-0.53299	0.078095	0.236421	0.270813	0.283818	0.589319	-0.73737	0.057449	-0.31277	-0.73557	-0
34	4	×OD4×He shot wide of the goal mouth.	-0.61174	0.173425	0.347637	-0.15977	-1.19368	0.095003	0.243637	-0.02827	-0.03994	0.670254	0.078198	-0.01665	-0.23078	-0.1218	0.102638	0.223839	0.450901	0.031692	-0.11488	-0.81428	0.302515	-0.58971	-0.42955	-0
35	4	×OD4×She wiped the mouth of the bottle before drinking	-0.78439	0.069043	0.253768	-0.19674	-0.64354	0.021513	0.096655	0.349658	0.027349	0.985118	-0.08074	0.079709	-0.8138	-0.27262	0.579832	0.439477	0.621638	-0.38895	0.49247	-0.9182	0.24061	-0.00336	-1.10505	-0
36	5	×OD5×He has a foul mouth on him!	-0.75276	-0.40415	-0.1048	-0.36122	-0.24003	0.148526	0.078271	0.110585	-0.07184	0.741633	0.125609	-0.22383	-0.30661	-0.34412	0.28982	-0.01721	0.084448	-0.2256	0.237457	-0.55174	0.031438	-0.24545	-0.48135	0.1
37	5	×OD5×Watch your mouth!	-1.40607	-0.29923	-0.39724	0.169193	-0.29609	0.112033	0.019601	0.314089	0.139288	0.877292	-0.01679	-0.08178	-0.20045	-0.12814	0.561012	-0.01592	-0.02521	-0.10711	0.255632	-0.41119	0.122189	0.007646	-0.47064	-0
38	6	×BNC×But there was humour in her tone, and a sparkle in	-0.77747	-0.17997	-0.50207	-0.7862	-0.02643	-0.19891	-0.48724	0.141441	-0.04051	1.030163	-0.63458	-0.22408	-0.2181	-0.32266	0.819154	0.126482	-0.52263	-0.22511	0.510833	-0.90022	-0.46423	-0.37246	-0.55891	0.0

mouth-OD-BNC-utf8-embeds

Methods | Dimension reduction 1024D→2D

- t-SNE (van der Maaten & Hinton, 2008) is a non-linear method that constructs a probability distribution over pairs of high-dimensional data points and a similar distribution over pairs of low-dimensional points, and it minimizes the difference between these two distributions using gradient descent in an iterative fashion. t-SNE is considered very effective at preserving the local structure of data at the expense of non-local structure.
- Isomap (Tenenbaum, de Silva & Langford, 2000) uses geodesic distance, which is a path between two points on a surface – rather than along a straight line. A graph is created by connecting neighbouring points and computing the geodesic distance between each pair of points.
- Spectral clustering: the top eigenvectors of the Laplacian matrix are considered to capture the global structure of the data.
- MDS creates a low-dimensional representation by minimizing the difference between distances of data point pairs in the high-dimensional space and pairwise distances in the low-dimensional space.

Methods | Observation of 1024D & 2D diagrams

- visual observation, looking for clusters, patterns
- *k*-means clustering (5000 iterations, 20 reruns)

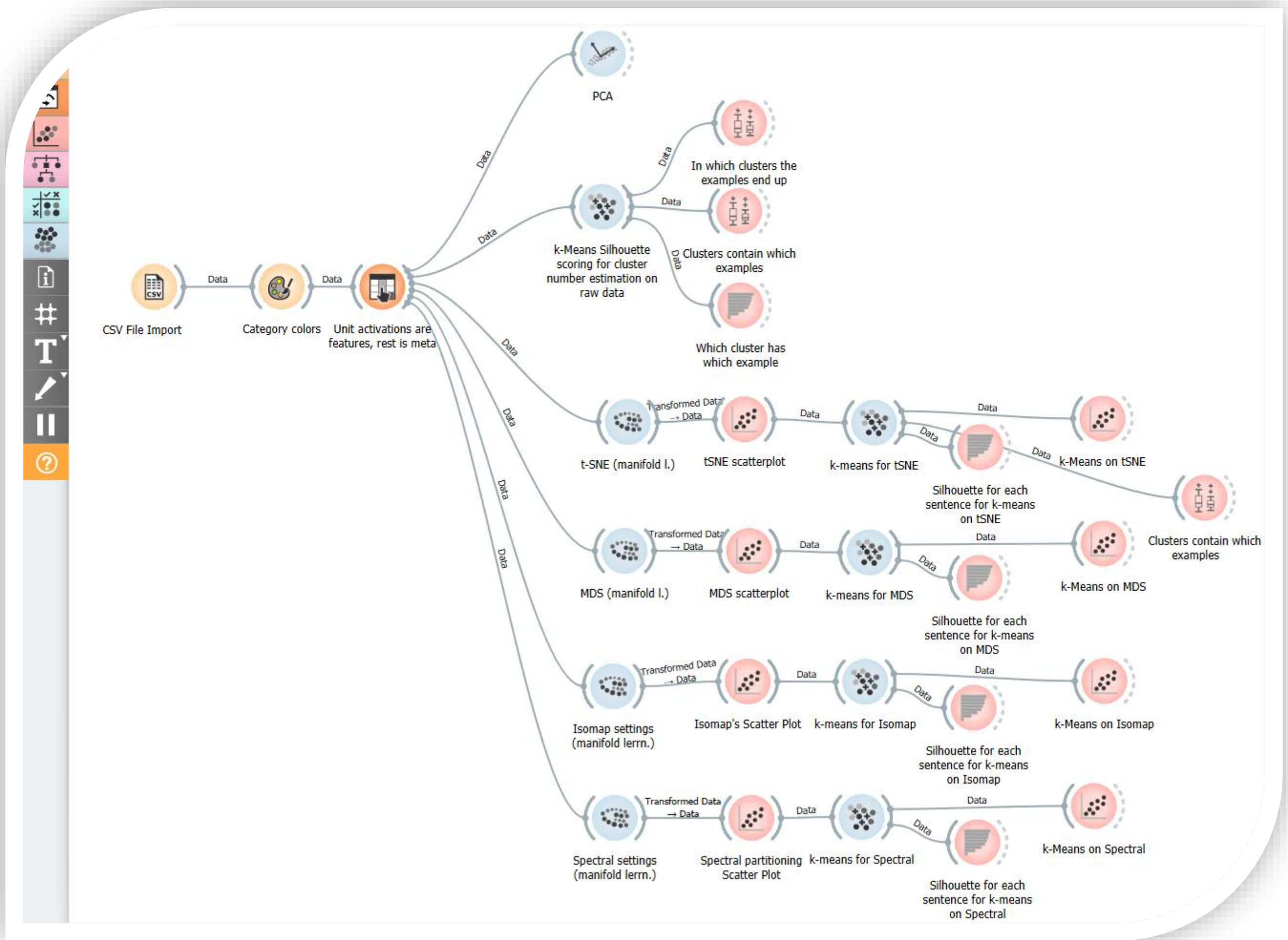
  For selecting *k*: Silhouette scoring (Rousseeuw, 1987); a measure of how well data points fit into their clusters, and it “shows which objects lie well within their cluster, and which ones are merely somewhere in between clusters” (ibid.). A higher score indicates better clustering.

UNIVERSITY of DEBRECEN

T·U·D·A
UNIVERSITY OF DEBRECEN
FACULTY OF HUMANITIES
SCHOLARLY FUND

Methods | The software tool we used

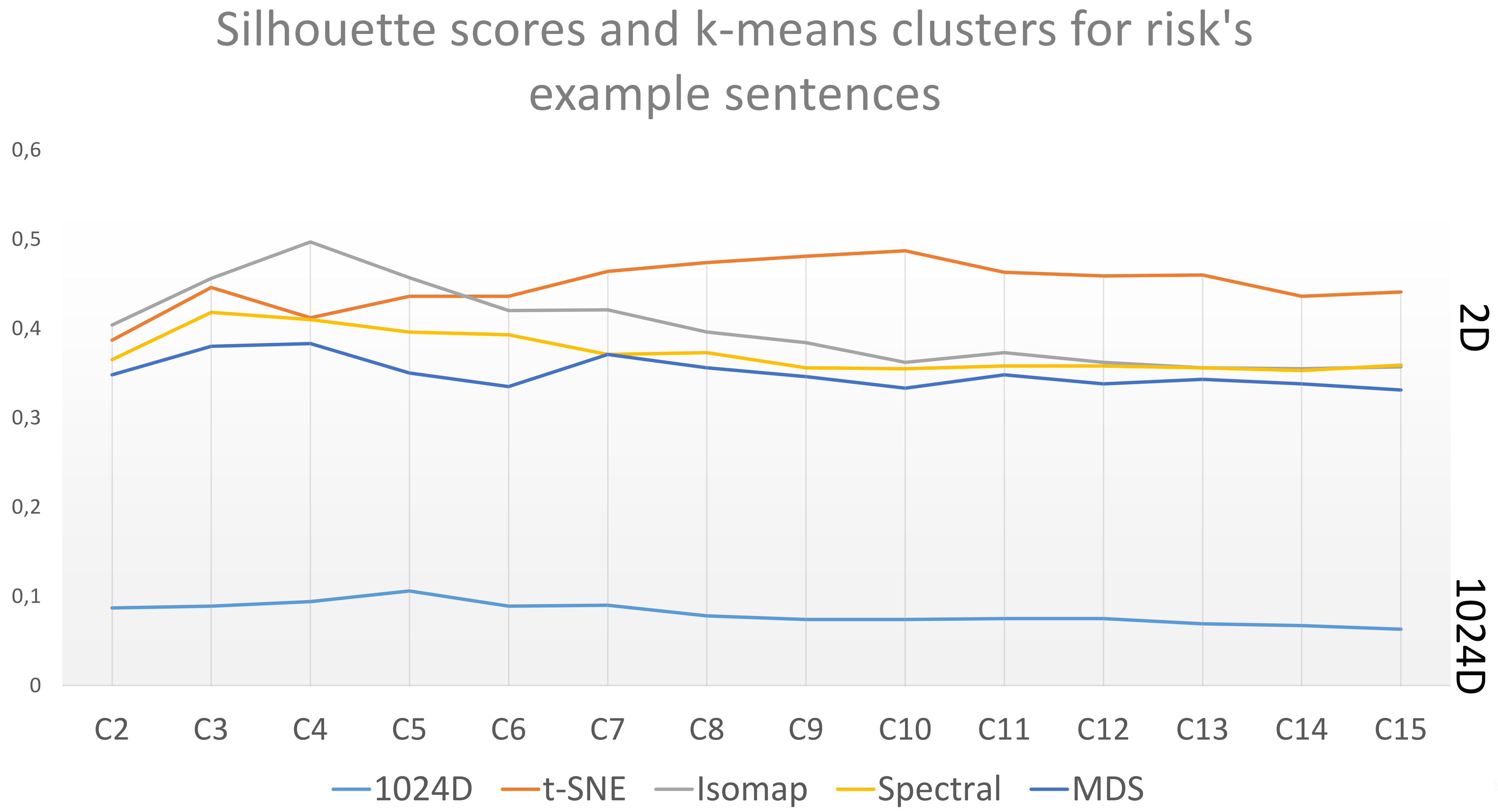
The software used for dimension reduction and *k*-means clustering, also the source of our illustrations: *Orange Data Mining toolkit* (Demsar et al., 2013; https://orangedatamining.com)



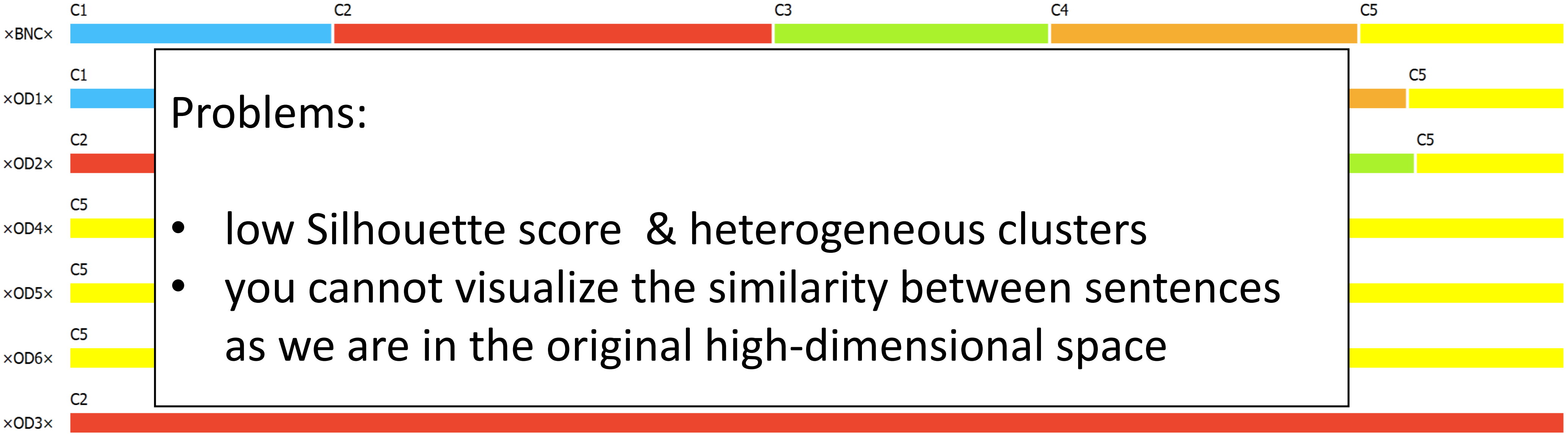
UNIVERSITY of DEBRECEN

T·U·D·A
UNIVERSITY OF DEBRECEN FACULTY OF HUMANITIES SCHOLARLY FUND

Results | Silhouette scores for clustering *risk*



Results | *risk* in BERT's original 1024D vector space



Problems:

- low Silhouette score & heterogeneous clusters
- you cannot visualize the similarity between sentences as we are in the original high-dimensional space

The verbal senses of *risk* clustered together in C5, but it was r…

C1 included the nominal uses only, medical risks dominate… risks a…

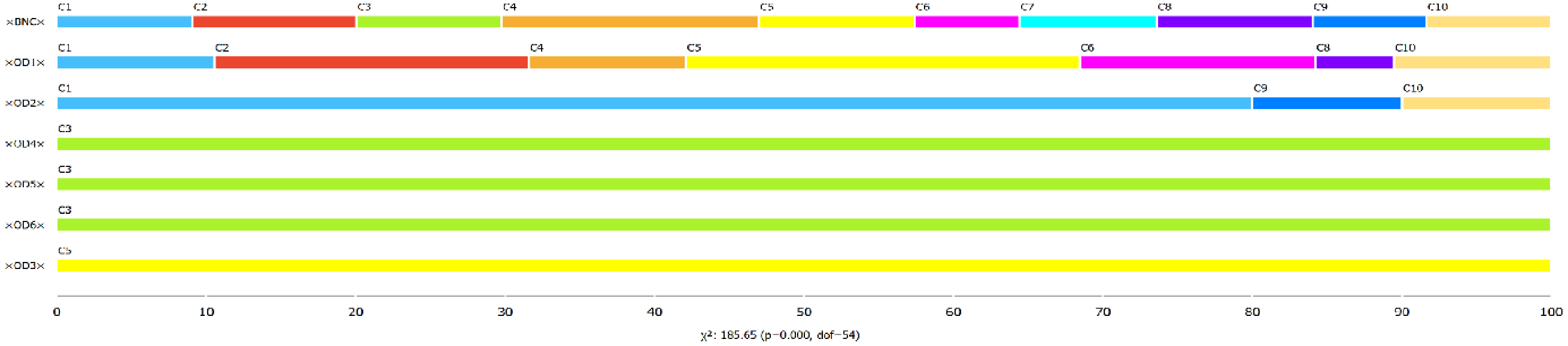
C2 was mostly associated with financial risks, with some health-related ones.

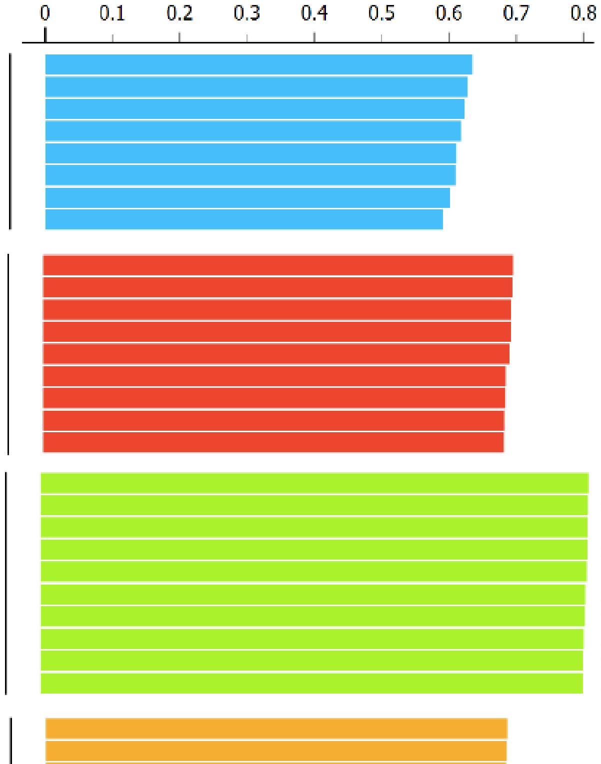
Sentences in C3 referred to social, environmental, economic and m…

Sentences in C4 generally referred to risky situations without specification and also associated *risk* with business loss and body injuries.

UNIVERSITY OF DEBRECEN FACULTY OF HUMANITIES SCHOLARLY FUND

Results | *k*-means clusters after t-SNE, k=10, *risk*





×BNC×'Generally that I had to be convinced any person posed a risk - and to give a warning before I fired.'
×BNC×Is the Minister satisfied that those humanitarian needs have been met and are being met, or that they can be met so long as there is a risk to the Kurdish popu
×BNC×Because of the increased risk due to the speeds of which these vessels are capable the clause restricts the cover granted substantially.
×BNC×It is the risk to public order inherent in the defendant's words or conduct that represents the harm struck at by the section.
×BNC×"Stone & Dobinson's test of"" obvious risk to health or welfare"" would broaden manslaughter perhaps unacceptably, and the position is now that accepted by t
×BNC×These are important areas, but the writers usually concentrate on the technical and physical aspects of securing computer systems against external threats whi
×BNC×Th
×BNC×Ye

×BNC×In
×BNC×The
×BNC×Clo
×BNC×Sm
×BNC×The
×BNC×Scr
×BNC×Evi
×BNC×Pat
×BNC×Hig

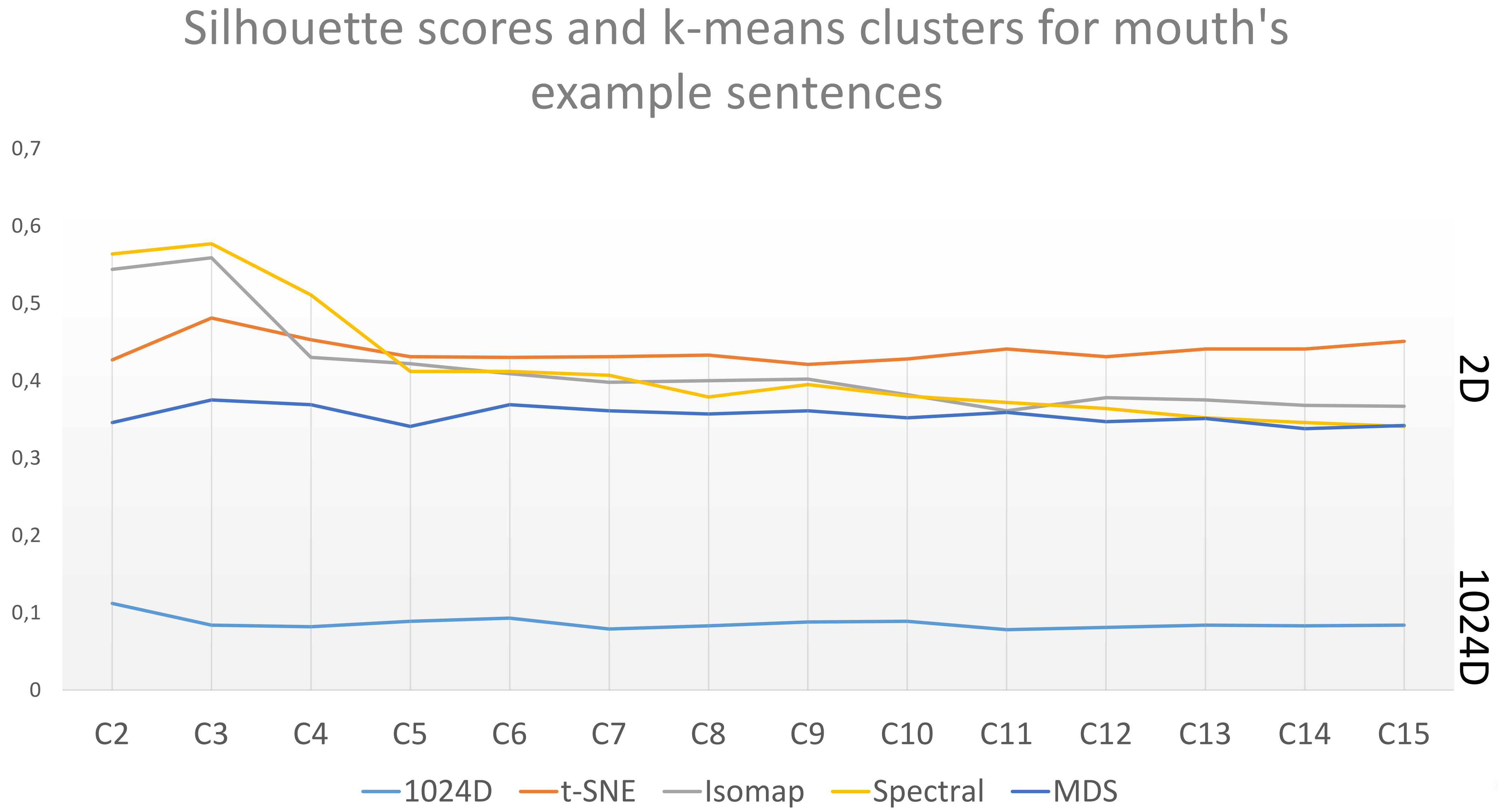
×BNC×One
×BNC×The
×BNC×'Yo
×BNC×Car
×OD6×We
×BNC×You
×BNC×Or
×BNC×Eve
×BNC×Can Bob Halton hear the protest and risk giving up some of his 'I love me'?
×BNC×In the current climate, few executives were prepared to risk airing their anxieties publicly, but privately they express a wide range of fears.

×BNC×Such a system calls for kites with similar stability characteristics, otherwise there is a risk of tangling the two kitelines, as the shared weight draws them togeth
×BNC×There is a real risk of it being regarded not as a mate but a meal.

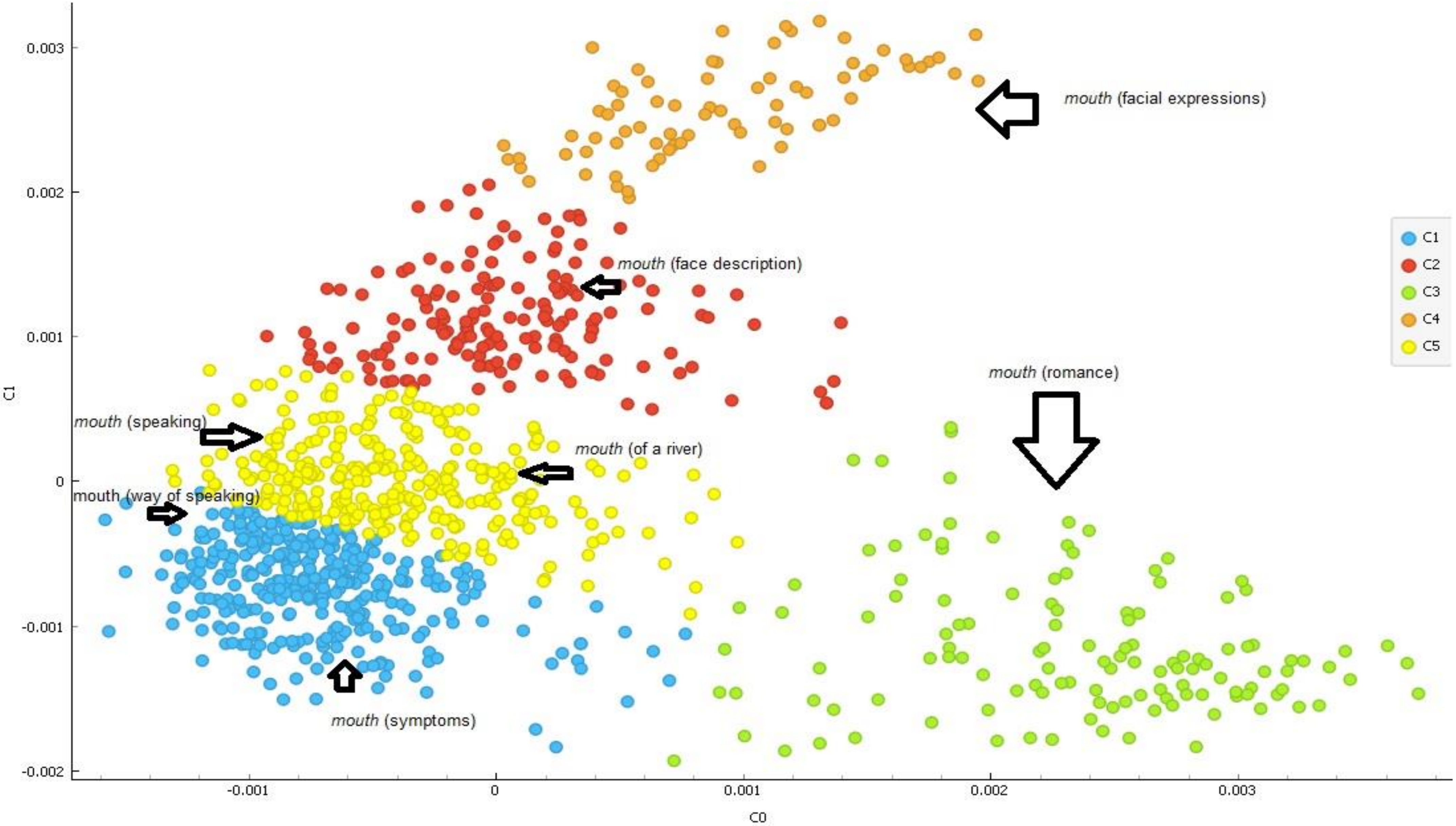
C1 included *be risk to* NP
C2: *increased/reduced/high/low risk of* NP
C4: *risk of -ing, risk+that+*clause
C6: Adj+*risk*+N
C7: *at risk*
Health-related risks: C2, financial risks: C5

UNIVERSITY OF DEBRECEN FACULTY OF HUMANITIES SCHOLARLY FUND

Results | Silhouette scores for clustering *mouth*

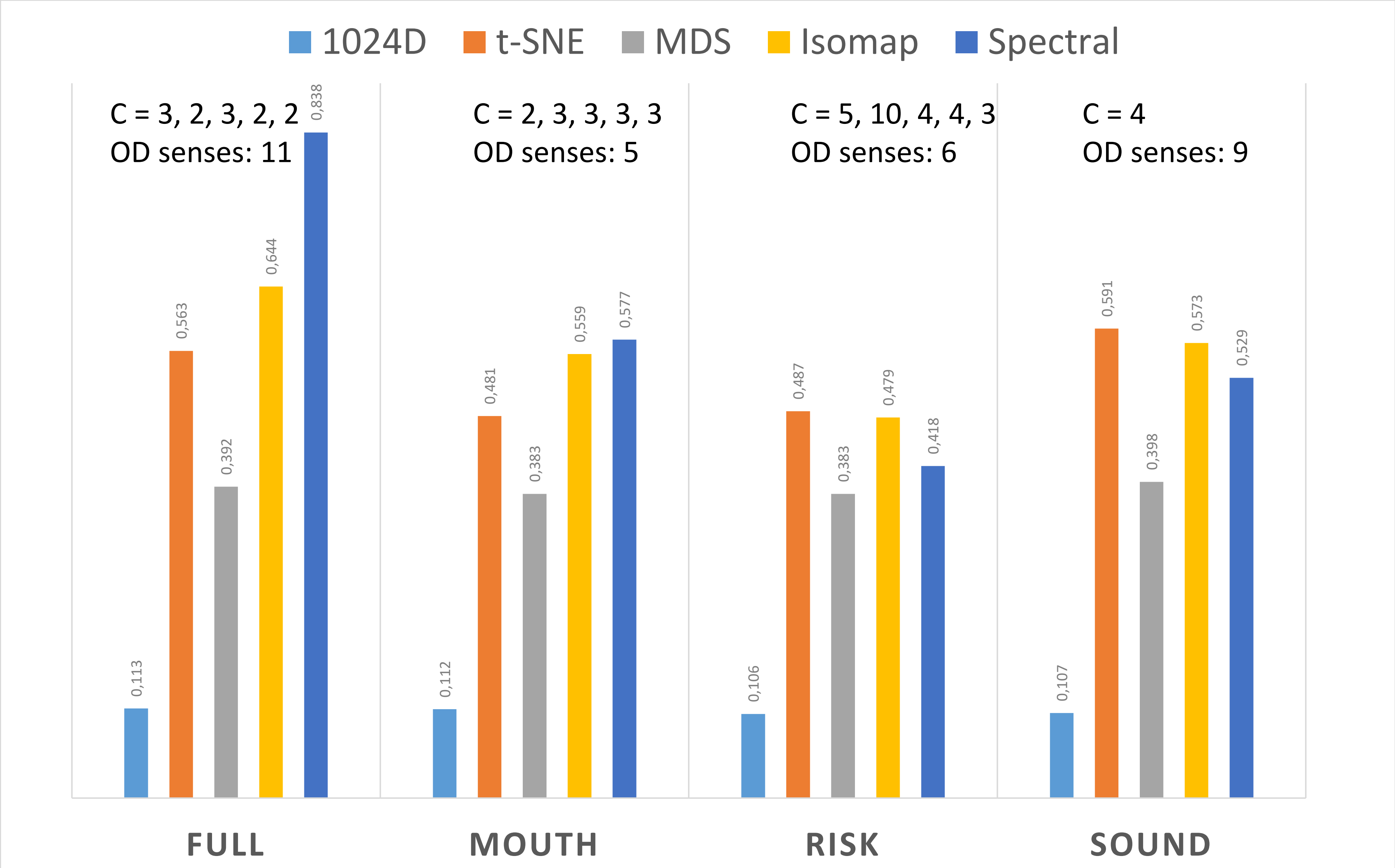


Results | *mouth* in Spectral visualization & clusters for *k*=5



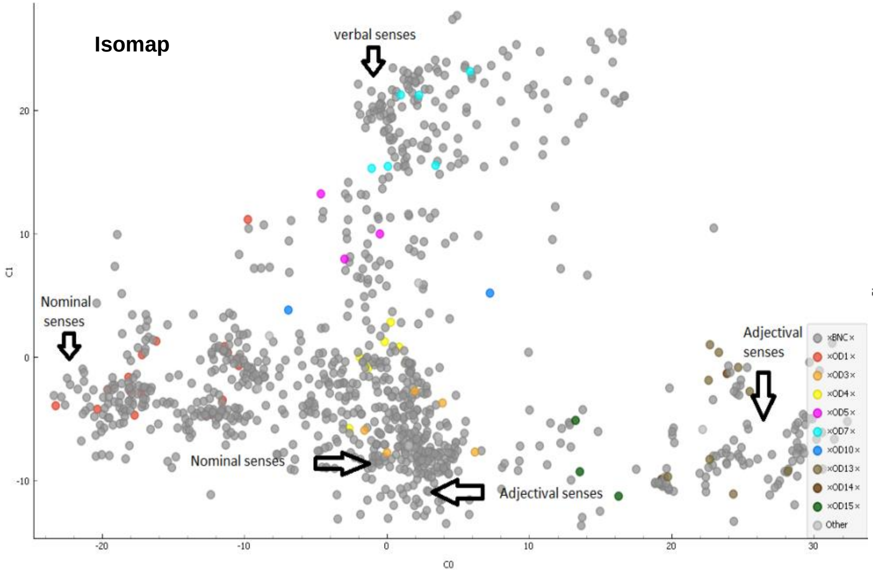
TUDA
UNIVERSITY OF DEBRECEN FACULTY OF HUMANITIES SCHOLARLY FUND

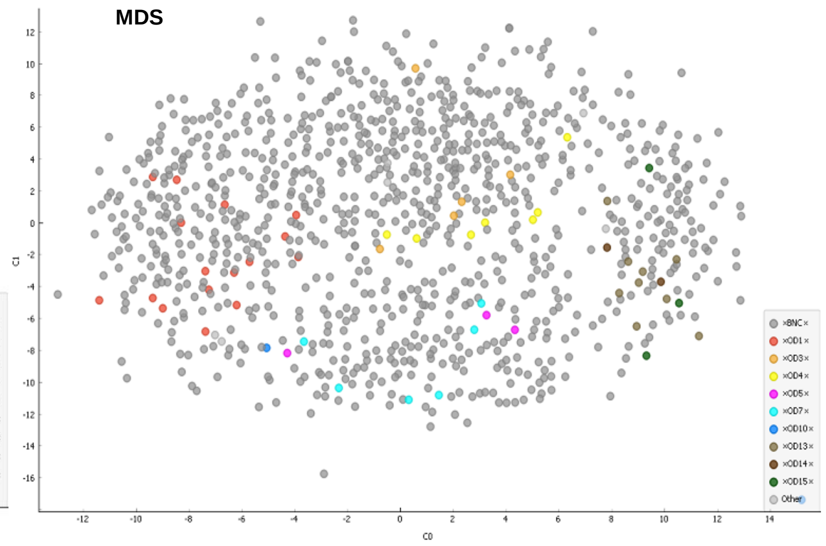
Results | The highest Silhouette scores for the four words before and after dimension reduction

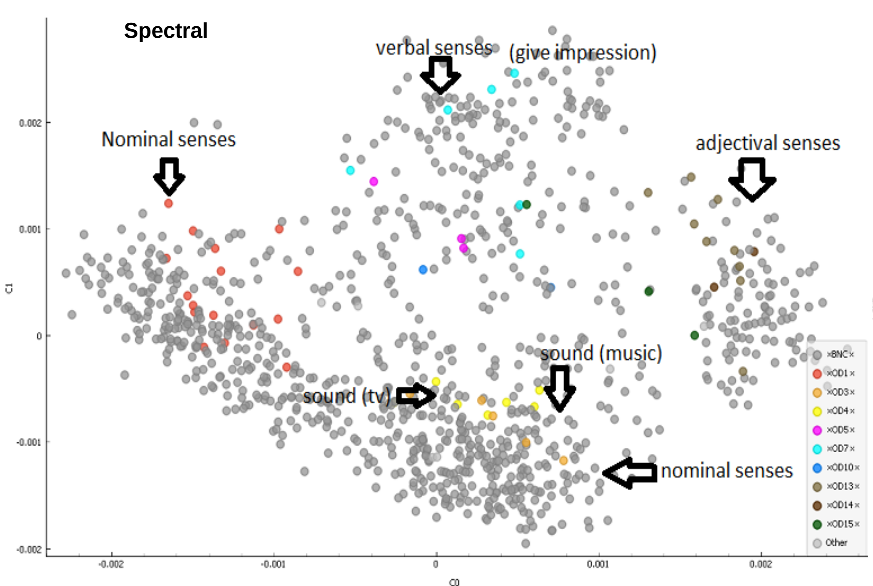


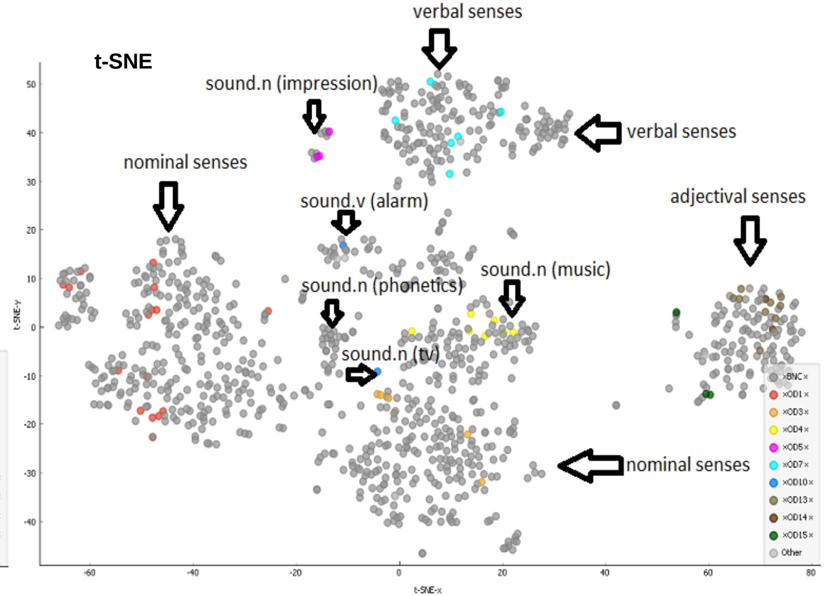
TUDA
UNIVERSITY OF DEBRECEN FACULTY OF HUMANITIES SCHOLARLY FUND

Results | *sound* in 4 visualizations (overview)

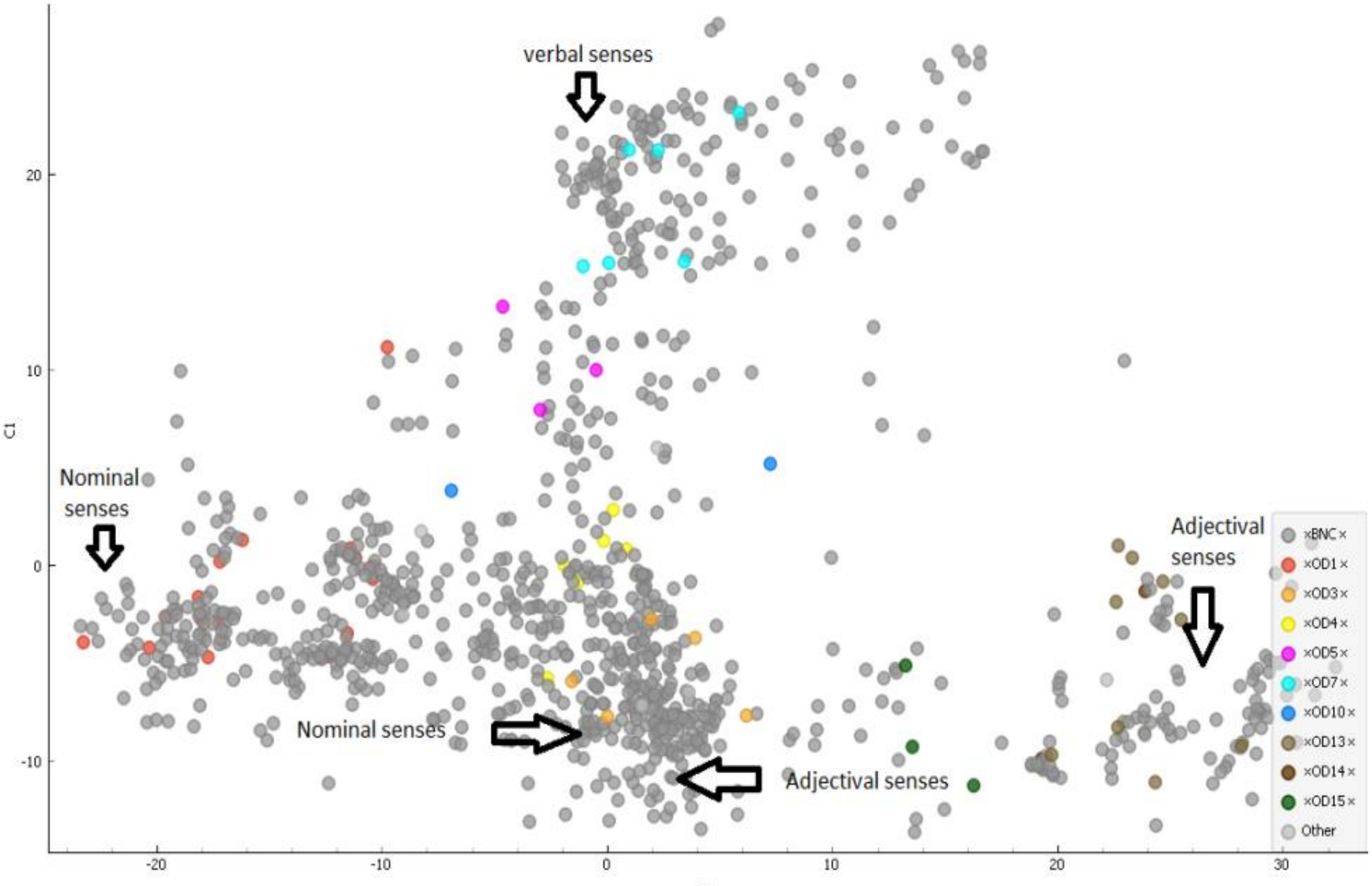




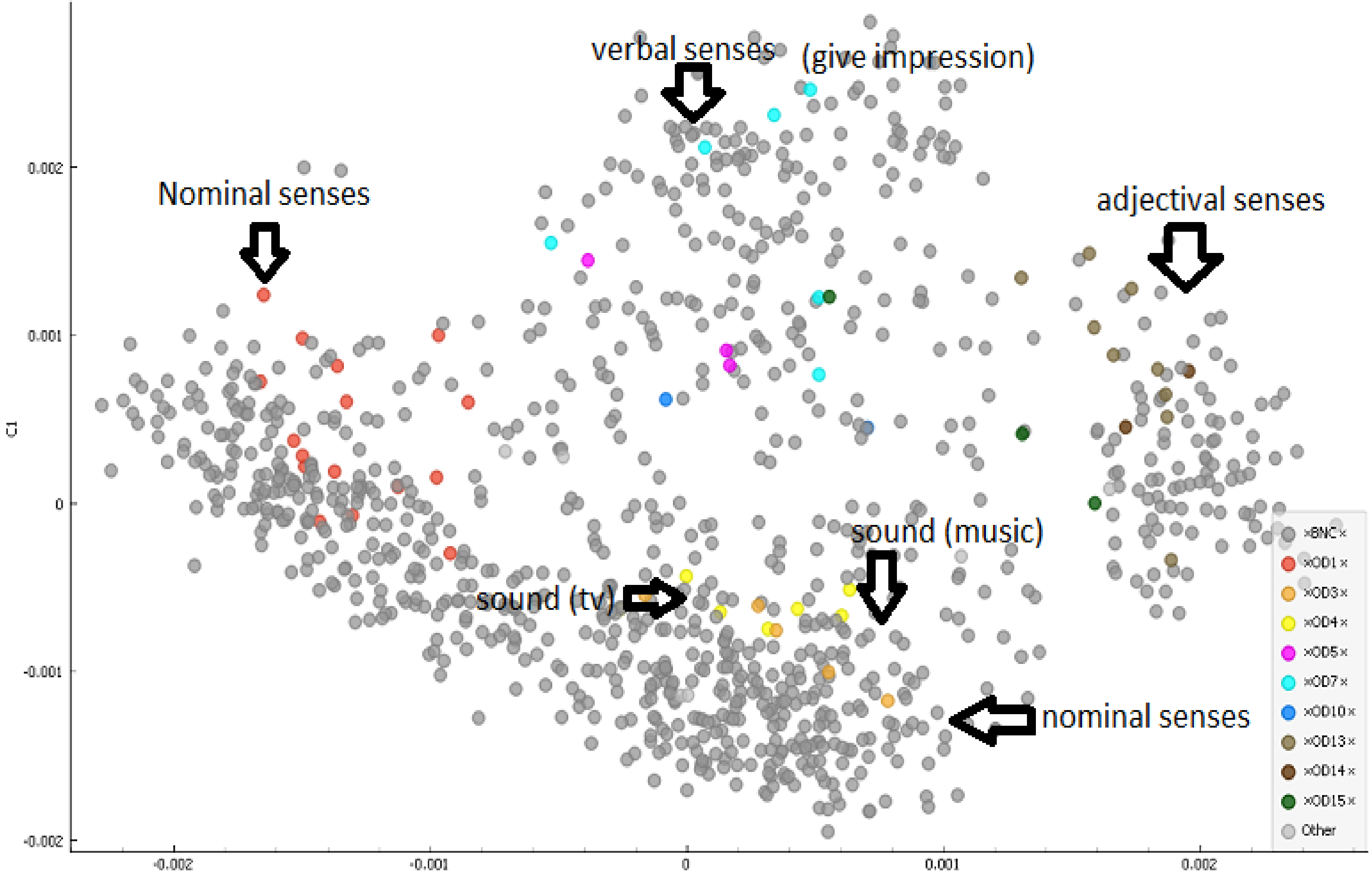




Results | *sound* in **Isomap**

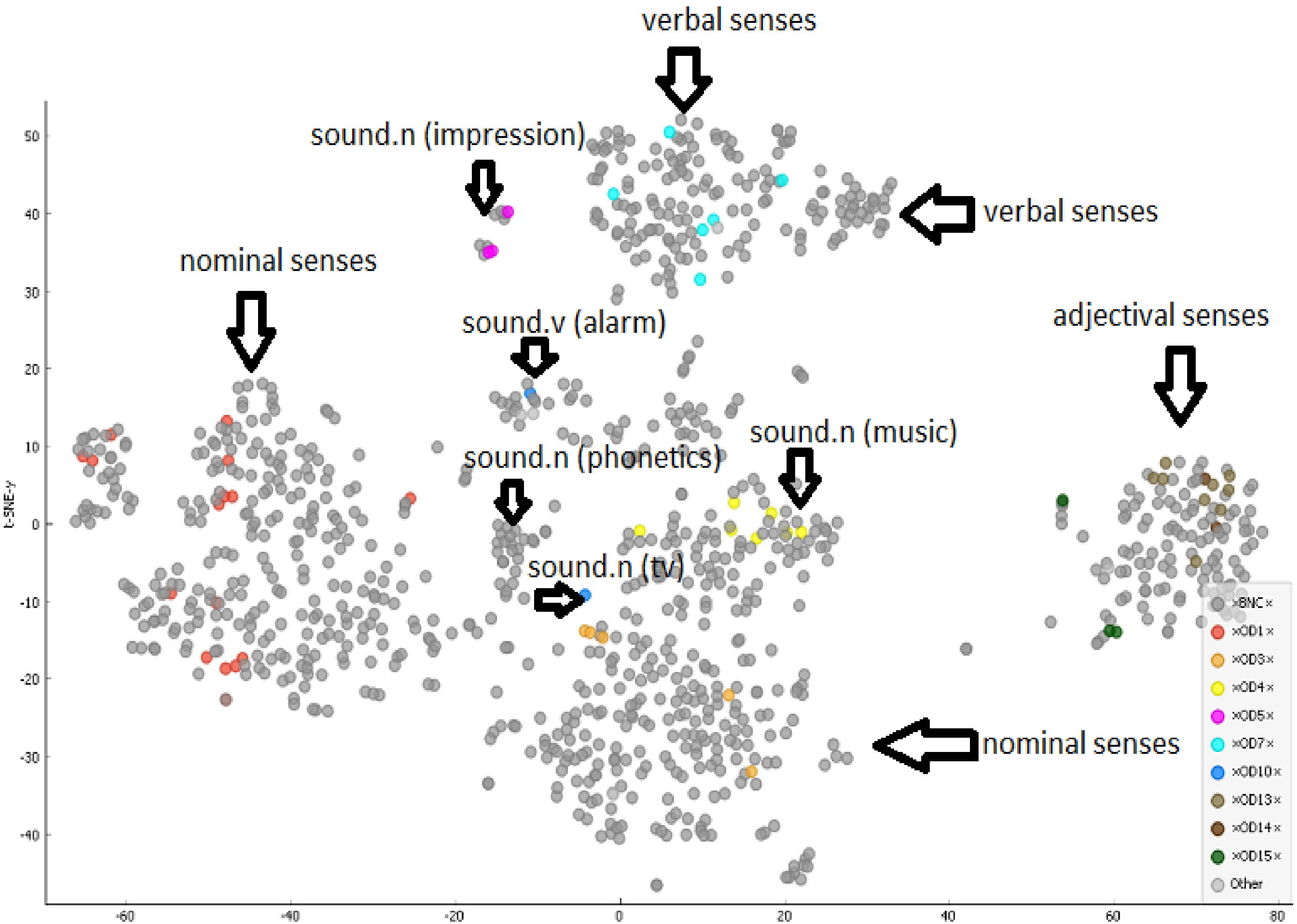


Results | *sound* in **Spectral**



UNIVERSITY OF DEBRECEN FACULTY OF HUMANITIES SCHOLARLY FUND

Results | *sound* in **t-SNE**



Parameter choices for the dimension reduction methods

Dimension reduction	Settings
t-SNE	perplexity = 20 (also tested: 10, 30) distance = Eucledian (also tested: Manhattan, Chebychev) initialization = PCA max. iterations = 3000 learning rate = 200
MDS	initialization = PCA max. iterations = 5000
Isomap	neighbours = 20
Spectral	affinity = RBF kernel (also tested: Nearest neighbour)

Discussion

- In our experiments, unsupervised separation between the metaphoric, metonymic and literal senses of words such as *mouth* and *sound,* based on the distributional features of the word uses, is reasonably good.
- The uses of words with relevance to specific semantic fields (e.g., *risk* in financial domains, *mouth* to make facial expressions, *full* with relevance to emotions) stood out in the automatically generated clusters.
- In almost all cases, Silhouette scoring for 2D representations recommended fewer categories than the number of Oxford Dictionary sense categories. Some dictionary distinctions were preserved within the sub-clusters (e.g., *sound* of music vs. *sound* of TV and radio), but others were lost (e.g. the four verbal senses of *risk*).

Conclusion

- The BERT-based, distributionally-motivated clusters did not correspond to the number of dictionary senses, but they did show BERT's sensitivity to semantic and syntactic similarities between word uses.
- Before dimension reduction, Silhouette scores of the *k*-means clusters were low, and so was the qualitative cohesion between the sentences in the cluster.
- Visualizing BERT representations in 2-dimensional spaces using Spectral, t-SNE and Isomap showed quantitative and qualitative improvements that can be beneficial to lexicographers. Not only the Silhouette scores of the clusters increased, but also semantic and syntactic similarities appeared in the clusters.

UNIVERSITY of DEBRECEN

T·U·D·A
UNIVERSITY OF DEBRECEN
FACULTY OF HUMANITIES
SCHOLARLY FUND

Conclusion

- MDS was inferior to the 3 remaining manifold learning algorithms in our case study.
- These visualizations can be helpful in enriching dictionary entries with additional, corpus-based examples; the closest BNC sentences to the dictionary examples mostly reflected very similar semantic and syntactic patterns.
- In our charts, we also saw thematically-motivated clusters of BNC sentences that were ignored during exemplification of the OD headword (e.g., the uses of the word *mouth* in romantic literature).

Acknowledgement

This publication was supported by the *University of Debrecen Faculty of Humanities Scholarly Fund*.

References

- Demsar J., Curk, T., Erjavec, A., Gorup, C., Hocevar, T., Milutinovic, M., Mozina, M., Polajnar, M., Toplak, M., Staric, A., Stajdohar, M., Umek, L., Zagar, L., Zbontar, J., Zitnik, M., Zupan, B. (2013.) Orange: Data Mining Toolbox in Python. *Journal of Machine Learning Research*, 14, pp. 2349−2353.
- Harris, Z. S. (1954) Distributional Structure. *Word*, 10:2-3, pp. 146–162, DOI: 10.1080/00437956.1954.11659520
- Rousseeuw, P. J. (1987). Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. *Journal of Computational and Applied Mathematics*, 20, pp. 53–65.
- Tenenbaum, J. B., de Silva, V. & Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science,* 290(5500), pp. 2319–2323.
- van der Maaten, L. & Hinton, G. (2008). Visualizing Data Using t-SNE. *Journal of Machine Learning Research*, 1, pp. 1–48.