**Ověřování textové koherence pomocí korpusu SYN2020**

Postupy:

* nejčastější kolokace některých výrazů
* fungování odkazovacích výrazů
1. **Nejčastější kolokace vybraných výrazů**

Výchozí slovo: *tento problém* – zevšeobecňující substantivum s identifikátorem tento

Zdůvodnění volby:

Substantivum *problém* má široký odkazovací potenciál, může odkazovat i k textu

Pojí se s řadou dalších slov, vytváří kolokace.

Na soudržnost textu můžeme usuzovat z frekvence kolokátů

1. Zadáme jednoduchý (tj. ne pokročilý) dotaz s výchozím atributem lemma:

tento problém

1. Objeví se konkordance.
2. Využijeme nástroj Kolokace, volíme velký rozsah:



Výsledná tabulka kolokátů, uspořádaná podle Log Dice (pravděpodobnosti tvoření skutečné kolokace)



LogDice srovnává frekvence obou slov a frekvenci celého spojení sledovaných slov, tj. např. tento *problém* a *vyřešit*, tento problém a poradit apod. Čím vyšší má hodnotu, tím je pravděpodobnější, že slova tvoří spojení, resp. tím častěji se vyskytují jako spojení v korpusu.

To lze ověřit kliknutím na „p“ u „p/n“ (levý sloupec): zjistíme, že nejvíce frekventované kolokáty skutečně tvoří většinou spojení s výrazem *tento problém*.

**Jak toto zjištění souvisí s koherencí:**

Frekventovaný výskyt slov, která se zvoleným výrazem (*tento problém*) často tvoří spojení, v jeho blízkém, nebo i vzdálenějším kontextu, svědčí o významovém propojení vět, jejichž součástí není zvolený výraz tento problém, s větou s tímto výrazem.

Věty, v nichž tvoří spojení výraz tento problém a sloveso potýkat se



Texty, v nichž jsou sloveso potýkat se a výraz tento problém v různých větách v blízkém kontextu, text je soudržný (koherentní).



Podobně viz *vyřešit* a *tento problém:* oba výrazy někdy tvoří přímo kolokaci, v jiných případech ne, přesto i tam je text soudržný, protože se výskyt těchto slov (řešit/vyřešit a tento problém) předpokládá a jejich výskyt v textové blízkosti je pokládán za projev koherence, i když přímo kolokaci netvoří.



1. fungování spojení tento problém jako odkazovacího výrazu

Můžeme ověřit jen prohlédnutím konkordancí:

* *tento problém* odkazuje na celou předchozí větu



* *tento problém* odkazuje k širšímu textu, minimálně ke dvěma předchozím větám

