

O‘ZBEK TILI FRAZEOLGIK BIRLIKLARIDA OMONIMIYANI N-GRAMM USULIDA ANIQLASH

Umid Yodgorov

Alisher Navoiy nomidagi Toshkent davlat
o‘zbek tili va adabiyoti universiteti,
Kompyuter lingvistikasi mutaxassisligi 1-bosqich magistranti

Annotasiya. Ushbu maqolada o‘zbek tili frazeologik birliklarida omonimiyani n-gramm usulida aniqlash usullari haqida fikr yuritiladi. Frazeologik omonimiyaning frazeologik birliklar orasidagi omonimiya hamda frazeologik birlik va erkin birikma orasidagi omonimiya mavjud. Korpusdagi n-gramm modellari esa shunday birliklarni aniqlash va ularni saralashga yordam beradi. Bunda chastotali tahlil yordam beradi. Har bir n-gram matnda necha marta uchrashi aniqlanadi. Turg‘un birliklar chastotasi tasodifiy n-gramlar chastotasiga nisbatan yuqori bo‘ladi.

Kalir so‘zlar: frazeologik birlik, omonimiya, n-gramm, chastotali tahlil.

Abstract. This article discusses the methods of identifying homonymy in phraseological units of the Uzbek language using the n-gram method. Phraseological homonymy includes homonymy between phraseological units and homonymy between a phraseological unit and a free combination. N-gram models in the corpus help to identify such units and sort them. Frequency analysis helps in this. It determines how many times each n-gram occurs in the text. The frequency of stable units is higher than the frequency of random n-grams.

Keywords: phrasema, homonymy, n-gram, frequency analysis.

Kirish

Ma’lumki, o‘zbek tilida frazeologik birliklarda omonimiya hodisasi uchraydi. Bir xil komponent (so‘z)lardan tuzilgan frazeologizmlar o‘zaro bog‘lanmagan (tamoman boshqa-boshqa) ma’no ifodalashi mumkin. Bu hodisa frazeologik

omonimiyadir. Masalan: *Elmurod uning gapini jiddiy, lekin og'irsinib tingladi. Keyin chug'urlashib uyni boshiga ko'targan bolalarga qayrildi. (P.T.) ...kun peshindan og'ganda, cho'l changini boshiga ko'tarib, qandaydir motosikl chiqib qoldi. (Sh. G'ulomov.) Siz xalqimizning baxt-saodati uchun kurashdingiz. Xalqimiz sizni hurmatlab, olqishlar bilan boshiga ko'taradi. (Oydin.) Asli meni xo'jayin o'rgatdi. Ilgari bir qatra og'zimga olmas edim. (O.) Komissiya hayron bo'lib qoldi, ammo hech kim meni qaytarishni og'ziga olmas edi. (A.Q.)*³⁶.

R.Sayfullayeva va b. ham frazeologik birliklar orasidagi omonimiyani xuddi shunday ta'riflashgan³⁷.

Frazeologik omonimiyaning bu turining quyidagi ko'rinishlari mavjud:

- 1) frazeologik birliklar orasidagi omonimiya;
- 2) frazeologik birlik va erkin birikma orasidagi omonimiya.

H.Jamolxonov frazeologik omonimiya hodisasini shunday izohlaydi: “Shaklan teng bo'lgan frazemalar tilshunoslikda omofrazemalar deb nomlanadi. Erkin birikma bilan frazema (turg'un birikma) o'rtasidagi shakliy tenglik hodisasi frazeologik omonimiya emas, frazeologik omonimlik sanaladi. Demak, bunday holatlarda «Omonimiya» va «omonimlik» atamaları o'zaro farqlanadi. Chunonchi:

1. *Ko'z yummoq (erkin birikma)* – «Ko'zini yumib olmoq» ma'nosida;
2. *Ko'z yummoq (turg'un birikma, frazema)* – «ko'ra-bila turib e'tiborsiz qoldirmoq» ma'nosida;

36 Шоабдурахмонов Ш., Асқарова М., Ҳожиёв А., Расулов И., Дониёров Х. Ҳозирги ўзбек адабий тили. – Тошкент «Ўқитувчи», 1980. – 446 б. – Б. 138.; Ўзбек тили лексикологияси. – Тошкент: «ФАН» нашриёти, 1981. – 305 б.

37 Sayfullayeva R., Mengliyev B., Boqiyeva G., Qurbonova M., Yunusova Z. Hozirgi o'zbek adabiy tili. O'quv qo'llanma. – T., “Fan va texnologiya”, 2009. – 416 b. – B. 133.

3. *Ko'z yummoq (turg'un birikma, frazema)* – «vafot etmoq» ma'nosida. Keltirilgan misollarning birinchi va ikkinchisida erkin va turg'un birikmalar, ikkinchi va uchinchisida esa turg'un birikmalar (frazemalar) o'rtasida shakliy tenglik borligini ko'ramiz. Birinchi va ikkinchi misollardagi shakliy tenglik omonimlik sirasiga, ikkinchi va uchinchi misollardagi shakliy tenglik esa omonimiya sirasiga kiradi³⁸.

Bundan ko'rinadiki, o'zbek tilidagi iboralar qidiruvi natijasida frazeologik birliklar va erkin birikma orasida omonimiya masalasi yechilgan bo'lishi kerak. Bu masalani yechishda n-gram modellarini qo'llash mumkin. Quyida o'zbek tili matnlarida frazeologik birlik va erkin birikma orasidagi omonimiyani yechishning n-gram usullari xususida fikr yuritamiz.

Asosiy qism

n-gram usuli haqida

n-gram usuli turg'un lisoniy birliklarni izlash va tahlil qilish uchun optimal vosita bo'lib xizmat qilishi mumkin. Biroq aniq natijaga erishish va korpusda bunday birliklar qidiruv natijasini ko'rish aniqligini oshirish uchun boshqa usul va yondashuvlarni birgalikda qo'llash muhim. Frazeologik birliklar bazasini yaratish uchun lingvistik korpusdan foydalanish nafaqat tildagi frazemalarning yorqin namunalarini aniqlaydigan, balki til o'qitish va tadqiqot uchun foydali resurs yaratadigan kuchli va zamonaviy yondashuvdir. Bu jarayon tizimli yondashuvni va frazeologik birliklarni ajratib olishdan tortib, ularning tuzilishi va tashkil etilishigacha bo'lgan barcha bosqichlarga rioya qilishni talab qiladi.

n-gram usuli turg'un lisoniy birliklar, masalan, frazeologik birliklar, idioma va aforizmlar, tilda tez-tez qo'llaniladigan boshqa iboralarni topishda juda samarali

38 Jamolxonov H. Hozirgi o'zbek adabiy tili. – Toshkent: “O'zbekiston milliy ensiklopediyasi” Davlat ilmiy nashriyoti, 2013. – 403 b. – B. 337.

vositadir. Quyida n-gram yordamida turg'un lisoniy birliklarni aniqlashning bir necha yondashuv va usullari keltiriladi:

1. n-gram hosil qilish

n-gram yaratish: matnni n-gramga bo'ling, bu yerda n - har bir ifodadagi so'zlar soni. Ungramlar (n=1), bigramlar (n=2), trigramlar (n=3) va uzunroq ketma-ketliklardan foydalanish mumkin.

```
from nltk import ngrams
```

```
matn = "Bahorda daraxtlar gullaydi."
```

```
n = 2 # gapda nechta so'z mavjud
```

```
bigramlar = list(ngrams(text.split(), n))
```

```
print(bigrams)
```

2. Chastota tahlili

Chastotalarni sanash. Har bir n-gram matnda necha marta uchrashi aniqlanadi. Turg'un birliklar chastotasi tasodifiy n-gramlar chastotasiga nisbatan yuqori bo'ladi.

```
from collections import Counter
```

```
bigram_freq = Counter(bigrams)
```

```
print(bigram_freq.most_common(10))
```

3. n-gramlarning tipik uzunligini aniqlash

Uzunlik tahlili: turli uzunlikdagi n-gramlar ko'rib chiqiladi. Ba'zan turg'un iboralar qisqa yoki uzun bo'lishi mumkin. Bunday hollarda ularni bigram yoki trigramlar sifatida tahlil qilish mantiqiy to'g'riroq.

4. Turg'unlik va kontekst

Kontekstual tahlil: n-gramlar statik ifodalar ekanligini aniqlash uchun kontekstdan foydalaniladi. Bunda kontekstdagi so'zlar o'rtasidagi o'xshashlikni o'lchash uchun so'z vektorini ko'rsatish modellaridan (Word2Vec, GloVe) foydalanish maqsadga muvofiq.

5. Filtrlash

Chastota chegarasini sozlash: turg'un iboralarni ifodalamasligi mumkin bo'lgan – noyob n-gramlarni istisno qilish uchun chegara o'rnatiladi. Masalan, korpusdagi birikmalarning minimal sonini sozlash mumkin.

6. Klasterlash

Klasterlash: n-gramlar yaratilib, ularning chastotasi hisoblangach, o'xshash ifodalar guruhlarini aniqlash uchun klasterlash (masalan, K-o'rtacha yoki iyerarxik klasterlash yordamida) amalga oshirilishi mumkin.

7. Kontekstni qayta ishlash

Kontekstga asoslangan modellar: kontekstga asoslangan kuchli ifodalarni aniqlash uchun BERT, GPT kabi chuqur o'rganish modellaridan foydalaniladi. Ushbu modellar oddiy n-gramlar tahlilida aniqlash qiyin bo'lgan frazeologik birliklarni aniqlashga yordam beradi.

8. Qo'lda tekshirish va ekspertiza

Aniqlangan turg'un lisoniy birliklarning to'g'riligi va yaxlitligini tasdiqlash uchun qo'lda tahlil qilish yoki tilshunoslar tomonidan tekshirilishi talab etiladi.

9. Turli misollarda qo'llash

Matn tahlilida qo'llanilishi: aniqlangan turg'un lisoniy birliklar sentiment tahlil, avtomatik tarjima qilish, ma'lumot olish va boshqa vazifalarda qo'llanadi.

Korpusdagi frazeologik birliklarni aniqlashda n-gramdan foydalanish chastota tahlili asosida turg'un iboralarni aniqlash imkonini beruvchi kuchli va samarali usuldir. Ammo shuni esda tutish kerakki, n-gram modellari ko'pincha frazeologik birliklarni ajratib olishning aniqligini oshirish uchun qo'shimcha ishlov berish va tahlil qilishni talab qiladi.

O'zbek tili korpusidan n-gram usuli bilan quyidagi frazeologik birlik paradigmasi aniqlandi:

og'zini ochib	og'zini ochgancha	og'zini ochmay
og'zi ochilgach	og'zini ochadi	og'zini ochtiraylik
og'zim ochilib	og'zingizni ochib	og'zin ochqilu
og'zingni och	og'zini ochishga	og'ziday ochilgan
og'zimni ochib	og'zingizni ochishga	og'zidin ochti
og'zini ochayotgani	og'zi ochiqmish	og'zim ochildi
og'zi ochilgan	og'zi ochuq	og'zimizni ochmasdan
og'zi ochiq	og'zini oching-chi	og'zingni ochib
og'zini ochadiganlardan	og'zi ochilib-yumildi	og'zini ochganicha
og'zi ochiqligicha	og'zin ochib	og'zing ochilib
og'zini ochmasligi	og'zini ochishidanmi	og'zini ochishi
og'zini ochsa	og'zin ochqay	og'zi ochilmay
og'zingiz ochiladi	og'zim ochib	og'zini ochmasdan
og'zini ochmasa	og'zini ochdi	og'zidek ochildi



og‘ziday ochib	og‘zini ochdi-yu	og‘zini ocholmaydi
og‘zimiz ochilib	og‘zini ochirib	

Xulosa sifatida aytish mumkinki, o‘zbek tili matnlarida frazeologik birliklar va ular bilan omonimiyani hosil qiluvchi so‘z birikmalarini n-gram modeli bilan avtomatik aniqlash mumkin. N-gram qidiruvi natijasida hosil bo‘lgan ro‘yxat qo‘lda tahlil qilinib, frazeologik birlik va erkin birikma qo‘lda aniqlanadi va matnda ularning har biriga tegishli teg biriktiriladi.

Adabiyotlar ro‘yxati

1. Шоабдурахмонов Ш., Асқарова М., Ҳожиёв А., Расулов И., Дониёров Х. Ҳозирги ўзбек адабий тили. – Тошкент «Ўқитувчи», 1980. – 446 б. – Б. 138.;
2. Ўзбек тили лексикологияси. – Тошкент: «ФАН» нашриёти, 1981. – 305 б.
3. Sayfullayeva R., Mengliyev B., Boqiyeva G., Qurbonova M., Yunusova Z. Hozirgi o‘zbek adabiy tili. O‘quv qo‘llanma. – T., “Fan va texnologiya”, 2009. – 416 b. – B. 133.
4. Jamolxonov H. Hozirgi o‘zbek adabiy tili. – Toshkent: “O‘zbekiston milliy ensiklopediyasi” Davlat ilmiy nashriyoti, 2013. – 403 б. – B. 337.