

UMR TEGTLASHDA UCHRAYDIGAN XATOLAR VA ULARNING OLDINI
OLISH YO‘LLARI

Xurmatoy Mullaboyeva

ToshDO‘TAU Kompyuter lingvistikasi mutaxassisligi 1-kurs magistranti

Annotatsiya: Mazkur maqolada Universal Meaning Representation (UMR) formatida semantik teglash jarayonida tez-tez yo‘l qo‘yiladigan xatolar tahlil qilinadi. Tadqiqot davomida o‘zbek tilidagi 200 dan ortiq gap UMR formatda qo‘lda teglanib, yuzaga kelgan muammolar tizimlashtirildi. Natijalar shuni ko‘rsatadiki, xatolarning asosiy qismi semantik rollar farqlanmasligi (masalan, goal va place, cause va purpose), graf yo‘nalishlarining noto‘g‘ri berilishi (mod va mod-of, place va place-of), hamda diskurs va pragmatik rollarning (topic, content, domain) chalkashib ketishi bilan bog‘liq. Maqolada har bir xato turli misollar orqali ko‘rsatilib, ularning oldini olish uchun metodik tavsiyalar beriladi.

Kalit so‘zlar: UMR, semantik teglash, annotatsiya xatolari, goal, place, cause, purpose, mod, mod-of, place of, topic, content, domain.

Abstract: This study examines common error patterns in semantic annotation using the Universal Meaning Representation (UMR) framework. A dataset of over 200 Uzbek sentences was manually annotated in UMR, and the resulting issues were systematically identified and categorized. The findings show that the majority of errors stem from (i) confusion between closely related semantic roles (e.g., goal vs. place; cause vs. purpose), (ii) incorrect assignment of graph directionality (e.g., mod vs. mod-of; place vs. place-of), and (iii) misapplication of discourse and pragmatic roles (e.g., topic, content, domain). The paper provides illustrative examples for each error type and proposes practical methodological guidelines to reduce annotation inconsistencies.

Keywords: UMR, semantic annotation, annotation errors, goal, place, cause, purpose, mod, mod-of, place-of, topic, content, domain.

1. KIRISH

So‘nggi yillarda semantik teglash tizimlari tabiiy tilni avtomatik qayta ishlash (NLP) sohasida muhim ahamiyat kasb etmoqda. Xususan, Universal Meaning Representation (UMR) formati turli tillar uchun yagona semantik tasvir yaratish imkonini beradi. Biroq amaliyot shuni ko‘rsatadiki, UMR ni qo‘llash jarayoni faqat texnik emas, balki kuchli lingvistik bilimni ham talab etadi.

Annotatorlar tomonidan yo‘l qo‘yilayotgan xatolar nafaqat individual graf sifatini pasaytiradi, balki mashina o‘rganish jarayonida noto‘g‘ri model shakllanishiga ham olib keladi. Shu sababli UMR annotatsiyasidagi tipik xatolarni aniqlash va tizimlashtirish dolzarb masala hisoblanadi.

Ushbu maqolaning maqsadi — UMR teglash jarayonida ko‘p uchraydigan xatolarni aniqlash, ularning sabablarini tahlil qilish va amaliy tavsiyalar ishlab chiqishdir.

2. Metodologiya

Tadqiqot uchun o‘zbek tilidagi 200 ta gapdan iborat korpus qo‘lda UMR formatda teglab chiqildi. Har bir annotatsiya jarayonida yuzaga kelgan shubhali yoki xato holatlar alohida qayd etildi.

Xatolar semantik rol, graf struktura va diskurs darajalariga ajratilib tahlil qilindi. Asosiy tahlil kategoriyalari quyidagilarni o‘z ichiga oldi:

Semantik rollar (*ARGO*, *GOAL*, *CAUSE*, *PURPOSE*, *RESULT*, *BENEFICIARY*)

Graf yo‘nalishi (*mod/mod-of*, *place/place-of*, *manner/manner-of*) (Banarescu et al., 2013)

Diskurs va pragmatik rollar (*TOPIC*, *CONTENT*, *DOMAIN*)

Aspekt va temporallik (*ASPECT*, *temporality*)

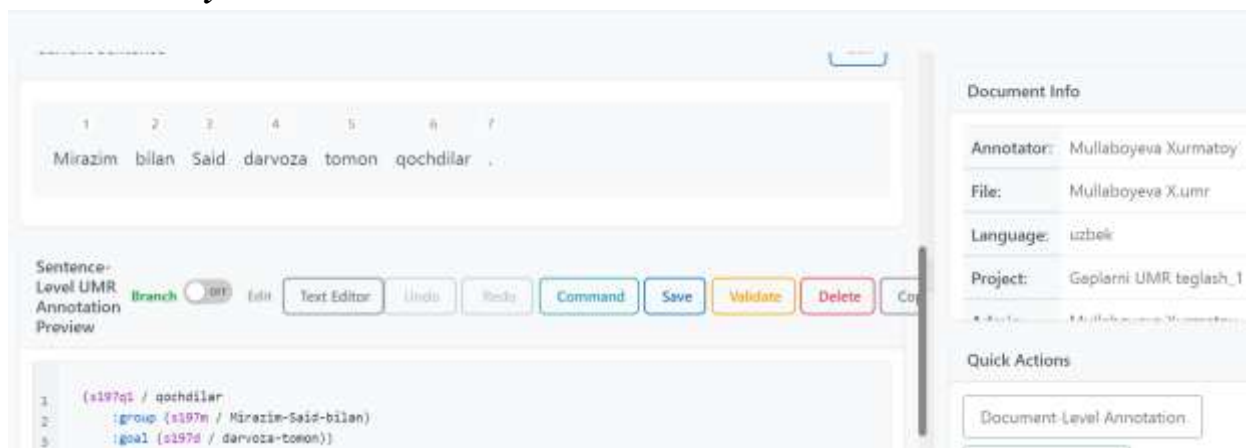
Har bir xato turi misollar asosida aniqlanib, qayta tahlil qilindi.

3. Natijalar

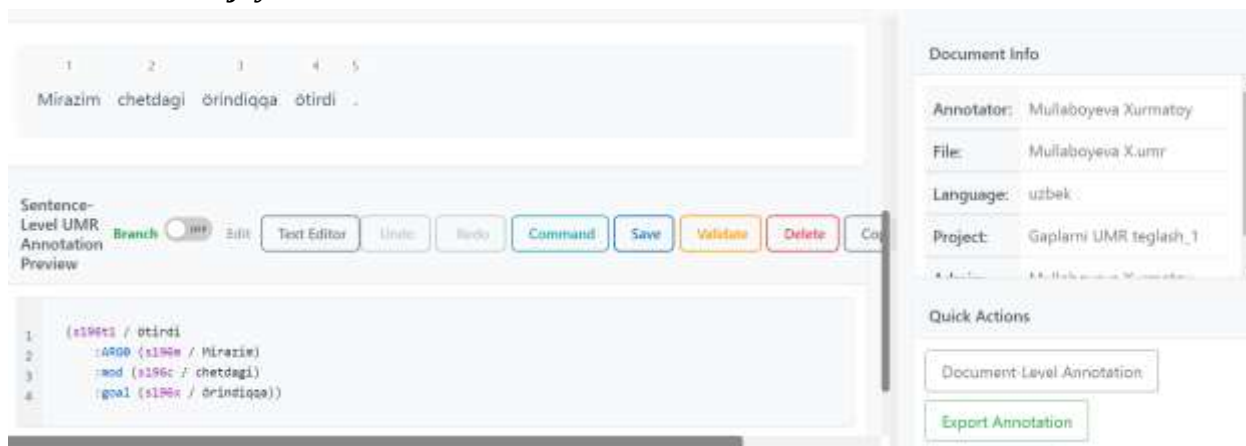
3.1. Semantik rollarning chalkashishi

Eng ko‘p uchraydigan xato — *goal* va *place* rollarining farqlanmasligi. Annotatorlar ko‘pincha “maktabga bordi” va “maktabda o‘qidi” kabi birikmalarda ikkisini bir xil teglaydi. Aslida birinchisi yo‘nalishni (*goal*), ikkinchisi esa joyni (*place*) bildiradi. (O’Gorman et al., 2020)

1-rasm. Goal-yo‘nalish.



2-rasm. Goal – joy.



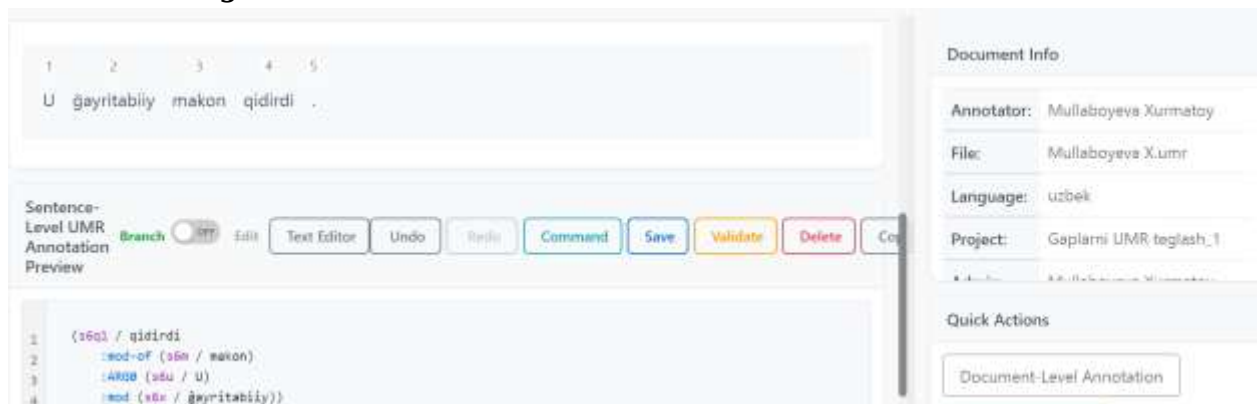
Yuqorida 1- rasmda darvoza tomon soʻzi *goal* qilib teglangan va bu gapda *goal* yoʻnalishni bildirmoqda. 2- rasmda oʻrindiqlik soʻzi *goal* qilib teglangan va bu gapda *goal* joy maʼnosini bildirmoqda.

Shuningdek, *cause* va *purpose* rollari ham tez-tez aralashtiriladi. “U kasal boʻldi sovuq sabab” konstruksiyasida *cause* boʻlsa, “U kasal boʻldi davolanish uchun” holatida *purpose* qoʻllanilishi kerak. (Jackendoff, 1990)

3.2. Graf yoʻnalishidagi xatolar

Modifikatorlar bilan bogʻliq xatolar ham keng tarqalgan. *mod* va *mod-of*, shuningdek *place* va *place-of* orasidagi farq koʻpincha tushunilmaydi. Aslida bu munosabatlar semantik jihatdan bir xil boʻlib, faqat *graf* yoʻnalishida farqlanadi. Annotatorlar esa baʼzan yoʻnalishni maʼno farqi sifatida qabul qiladilar. (Banarescu et al., 2013)

3-rasm. *mod* teglanish.



3.3. Diskurs va pragmatik rollarning notoʻgʻri ishlatilishi

Topic, content va domain rollari ham tez-tez chalkashadi. Masalan, “kitob haqida gapirdi” holatida topic boʻlishi kerak. “Kitobni oʻqidi”da esa *ARG1*. *Content* esa faqat ichki proposition mavjud boʻlganda qoʻllanadi (“u kelishini bildi”). (Jackendoff, 1990)

DOMAIN - predikatning qaysi soha / doira / konseptual maydon ichida amal qilishini bildiradi. Qaysi jihatdan? qaysi sohada? nimaga nisbatan? Bu joy ham emas, sabab ham emas - kontseptual kontekst. (Jackendoff, 1990)

DOMAIN tegi qoʻllaniladigan holatlar.

A) Baholash, fikr, sifatlash maʼlum sohada boʻlsa;

U matematika boʻyicha kuchli.

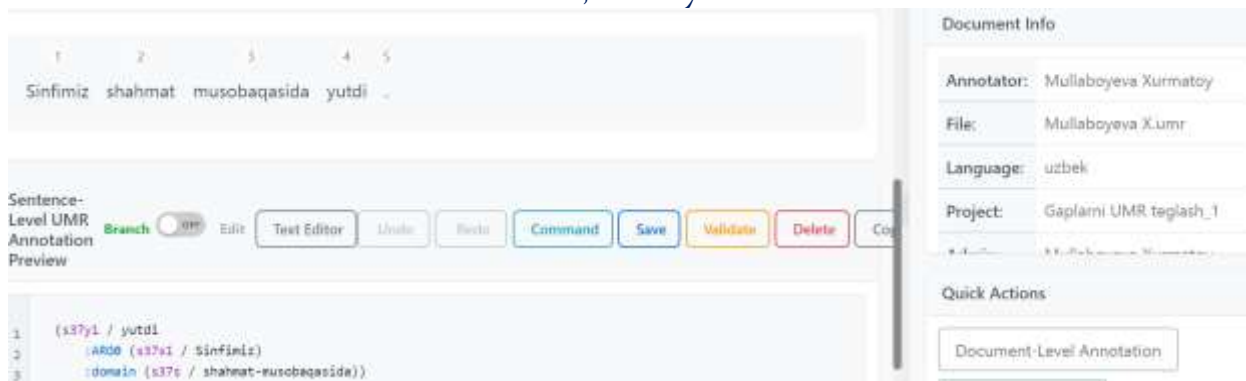
(s / kuchli

:*ARG0* (u / u)

:*DOMAIN* (m / matematika))

➡ kuchlilik qaysi jihatdan? → matematika

4-rasm. *UMR* teglashda domain tegi qoʻllanilgan gap.



Bu gapda ham sinfning qaysi sohada yutganligini aniqlab kelgani uchun shahmat musobaqasida *domain* qilib teglangan.

B) Emotsiya yoki holat muayyan obyektga nisbatan bo‘lsa;

U senga nisbatan sovuq.

Umr

(s / sovuq

:ARG0 (u / u)

:DOMAIN (s2 / sen))

C) Qobiliyat, bilim, munosabat sohada belgilanadi

Ali ingliz tilidan yaxshi.

(y / yaxshi

:ARG0 (a / Ali)

:DOMAIN (I / ingliz-til))

1-jadval. DOMAIN ≠ TOPIC ≠ GOAL

Teg	Savol	Farqi	Misollar
DOMAIN	qaysi jihatdan?	baholash doirasi	U sportda kuchli.
TOPIC	nima haqida?	nutq mavzusi	U sport haqida gapirdi.
GOAL	qayerga / kimga	yo‘nalish	U sportga bordi.

Quyidagi holatlarda *domain* qilib teglanmaydi;

1. Agar bu oddiy obyekt bo‘lsa → ARG1

2. Agar bu joy bo‘lsa → PLACE

3. Agar bu mavzu bo‘lsa → TOPIC

Agar olib tashlasak gap grammatik qolsa, lekin baholashning nimaga nisbatan ekani yo‘qolsa → DOMAIN

UMR teglashda *domain* predikatning amal qilish doirasini bildiradi va baholash, munosabat yoki qobiliyatning qaysi obyektga yoki soha ichida ekanini ko‘rsatadi.

DOMAIN — baholash / munosabat / holatning qaysi sohada ekanini bildiradi.

3.4. Aspekt va temporallikdagi muammolar

Annotatorlar ko‘pincha davomiy holatlarni (*durative*) oddiy harakatlar bilan bir xil ko‘rsatadilar yoki *temporality* ni umuman belgilamaydilar. Bu esa holatning vaqtinchalik yoki doimiy ekanini yo‘qotadi. (Artstein & Poesio, 2008)

Tadqiqot davomida *ASPECT* va *TEMPORALITY* teglari bilan bog‘liq xatolar ham muhim ulushni tashkil etgani aniqlandi. Annotatorlar ko‘pincha harakatning davomiyligi (*durative*), jarayon shakli (*progressive*) va natijaviy holati (*resultative*) o‘rtasidagi farqni yetarlicha ajratmaydi. Masalan, davomiy holatni bildiruvchi “yashaydi”, “kutib turardi” kabi fe‘llar ba‘zan oddiy hodisa sifatida belgilanib, ularga *ASPECT durative* tegi qo‘yilmaydi. Natijada harakatning vaqt bo‘yicha ichki tuzilishi yo‘qoladi.

Shuningdek, *temporality* bilan bog‘liq xatolar ham keng tarqalgan. Annotatorlar holatning vaqtinchalik yoki barqarorligini farqlashda qiynaladilar. Masalan, “Ali hozir kasal” kabi holatlar vaqtinchalik (*temporary*) bo‘lsa-da, ular ko‘pincha belgisiz qoldiriladi yoki barqaror xususiyat sifatida noto‘g‘ri talqin qilinadi. Bu esa holatning ontologik turini buzadi va keyingi avtomatik tahlil jarayonlariga salbiy ta‘sir ko‘rsatadi.

Bundan tashqari, *ASPECT* va *temporality* rollarining chalkashib ketishi ham kuzatiladi. Ba‘zi annotatsiyalarda davomiylikni ifodalash uchun *temporality* ishlatilgan yoki aksincha, vaqtinchalik holatlar *ASPECT* orqali ko‘rsatilgan. Aslida *ASPECT* hodisaning ichki vaqt tuzilishini, *temporality* esa holatning barqaror yoki vaqtinchaligini bildiradi.

4. Muhokama

Aniqlangan xatolar asosan uch sababga borib taqaladi: (1) semantik rollar orasidagi nazariy farqlar yetarlicha mustahkam o‘zlashtirilmaganligi, (2) *graf* struktura va semantik mazmun o‘rtasidagi farq tushunilmaganligi, (3) *diskurs* darajasi e‘tibordan chetda qolishi. (O’Gorman et al., 2020)

Quyida *graf* yo‘nalishi teglarining (*mod / mod-of*, *place / place-of*, *manner / manner-of*) o‘xshash va farqli tomonlari jadval ko‘rinishida, real misollar bilan berildi.

2-jadval. Graf yo‘nalishi teglarining solishtirma jadvali

Juftlik	O‘xshash tomoni	Farqi (asosiy)	Qac hon ishlatiladi	Tegl angan misol
<i>mod</i> / <i>mod-of</i>	Ikkalasi ham aniqlovchi munosabatini bildiradi	Yo‘nalis h: <i>mod</i> → bosh so‘zdan modifierni, <i>mod-of</i> → modifierdan bosh so‘zga	<i>mod</i> – obyekt markaz bo‘lsa; <i>mod-of</i> – aniqlovchi markaz bo‘lsa	(gul : <i>mod</i> chiroyli) vs (chiroyli : <i>mod-of</i> gul)
<i>place</i> / <i>place-of</i>	Ikkalasi ham joy munosabatini bildiradi	Yo‘nalis h: <i>place</i> → hodisadan	<i>place</i> – hodisa markaz;	(bor- 01 : <i>place</i> maktab) vs

		joyga, <i>place-of</i> → joydan hodisaga	<i>place-of</i> – joy markaz	(maktab : <i>place-of</i> bor-01)
manner / manner-of	Ikkalasi ham harakat uslubini bildiradi	Yo‘nalis h: <i>manner</i> → fe‘ldan uslubga, <i>manner-of</i> → uslubdan fe‘lga	<i>man</i> – harakat markaz; <i>manner-of</i> – uslub markaz	(yugur-01 : <i>manner</i> tez) vs (tez : <i>manner-of</i> yugur-01)

Ushbu juftliklarda semantik munosabat bir xil bo‘lib, farq faqat *graf* strukturaning yo‘nalishida namoyon bo‘ladi. Annotatsiyada odatda harakat yoki obyekt markaz bo‘lsa to‘g‘ri yo‘nalish (*mod*, *place*, *manner*) tanlanadi, aniqllovchi yoki kontekst markazga chiqarilganda esa teskari shakl (*mod-of*, *place-of*, *manner-of*) qo‘llanadi.

3-jadval. Graf yo‘nalishi teglarining solishtirma jadvali

	Teg turi	Gap	UMR teglangan ko‘rinish	Izoh
	mod	Oppoq mashina shitob bilan qo‘zg‘aldi.	(m / mashina : <i>mod</i> (o / oppoq))	— <i>oppoq mashinaning</i> aniqllovchisi
	place	Ali maktabda o‘qiydi.	(o / o‘qi-01 : <i>place</i> (m / maktab))	O‘qish hodisasi qayerda bo‘layotgani ko‘rsatilgan
	manner	Hasan chopib ichkariga kirdi.	(k / kir-01 : <i>manner</i> (c / chop-01))	Kirish harakati qanday usulda bo‘ldi
	mod-of	(Grafni teskari qurishda)	(o / oppoq : <i>mod-of</i> (m / mashina))	Aniqllovchi tugun markazga chiqarilgan
	place-of	(Grafni teskari qurishda)	(m / maktab : <i>place-of</i> (o / o‘qi-01))	Joy tuguni markaz, hodisa unga bog‘langan
	manner-of	(Grafni teskari qurishda)	(t / tez : <i>manner-of</i> (y / yugur-01))	Uslub tuguni markaz qilib olingan

ASPECT va *temporality* bilan bog‘liq xatolar asosan ikki sababga borib taqaladi. Birinchidan, annotatorlar grammatik zamon shakllari bilan semantik aspekt tushunchasini tenglashtirib yuboradilar. Masalan, “-yapti” shakli avtomatik ravishda progressive sifatida qabul qilinadi, holbuki ba’zi hollarda u umumiy holat yoki odatiy harakatni bildirishi

mumkin. Ikkinchidan, temporality tushunchasi ko‘pincha e‘tibordan chetda qoladi, chunki u ochiq grammatik markerlarga ega emas va kontekst asosida aniqlanadi. . (Banarescu et al., 2013)

4-jadval. ASPECT va Temporality teglari bo‘yicha xatolar (misollar bilan)

Gap	Noto‘g‘ri teglash	To‘g‘ri teglash	Xato turi / izoh
Ali hozir kasal.	(k / kasal)	(k / kasal :TEMPORALITY temporary)	Vaqtinchalik holat belgilanishi kerak edi
U uzoq kutdi.	(k / kut-01)	(k / kut-01 :ASPECT durative)	Davomiylik ko‘rsatilmagan
Hasan yugurib charchadi.	(y / yugur-01)	(y / yugur-01 :RESULT (c / charcha-01))	Natija alohida belgilanmagan
Qiz o‘qi-yapti.	(o / o‘qi-01 :ASPECT durative)	(o / o‘qi-01 :ASPECT progressive)	Progressive va durative chalkashgan
U anchadan beri yashaydi.	(y / yash-01)	(y / yash-01 :ASPECT durative)	Davomiy holat belgilanmagan
Ali doim sabrli.	(s / sabrli :TEMPORALITY temporary)	(s / sabrli)	Barqaror xususiyat vaqtinchalik qilib belgilangan
Bugun havo sovuq.	(s / sovuq)	(s / sovuq :TEMPORALITY temporary)	Kontekstga bog‘liq vaqtinchaliklik ko‘rsatilmagan
U tez-tez keldi.	(k / kel-01)	(k / kel-01 :ASPECT iterative)	Takroriylik belgilanmagan

Bu xatolar semantik modellashtirish sifatiga bevosita ta‘sir ko‘rsatadi. Davomiylik noto‘g‘ri belgilanganda, hodisalar orasidagi vaqt munosabatlari buziladi, temporality noto‘g‘ri berilganda esa barqaror xususiyatlar va vaqtinchalik holatlar aralashib ketadi. Natijada avtomatik xulosa chiqarish, sabab–natija aloqalarini aniqlash va voqealar ketma-ketligini modellashtirish jarayonlari noaniqlashadi.

5. XULOSA

UMR annotatsiyasi yuqori aniqlik va lingvistik sezgirlikni talab qiladi. Eng ko‘p uchraydigan xatolar — semantik rollarni chalkashtirish, graf yo‘nalishini noto‘g‘ri tushunish va diskurs rollarni noto‘g‘ri qo‘llash bilan bog‘liq. Ushbu maqolada keltirilgan tahlil va tavsiyalar annotatorlar tayyorlash jarayonini takomillashtirishga xizmat qilishi mumkin.

Shu sababli annotatorlarni tayyorlash jarayonida faqat texnik format emas, balki semantik nazariya ham chuqur o‘rgatilishi zarur. Ayniqsa, *ASPECT* va *TEMPORALITY* tushunchalari alohida, maxsus mashqlar orqali mustahkamlanishi lozim. Annotatorlar grammatik shaklga emas, semantik ma’noga tayangan holda qaror qabul qilishga o‘rgatilishi zarur.

FOYDALANILGAN ADABIYOTLAR:

1. Banarescu, L., Bonial, C., Cai, S., Georgescu, M., Griffitt, K., Hermjakob, U., Knight, K., Koehn, P., Palmer, M., & Schneider, N. (2013). Abstract Meaning Representation for Sembanking. Proceedings of the Linguistic Annotation Workshop, ACL.
2. Palmer, M., Gildea, D., & Kingsbury, P. (2005). The Proposition Bank: An Annotated Corpus of Semantic Roles. Computational Linguistics, 31(1), 71–106.
3. Pustejovsky, J. (1995). The Generative Lexicon. MIT Press.
4. Jackendoff, R. (1990). Semantic Structures. MIT Press.
5. Baker, C. F., Fillmore, C. J., & Lowe, J. B. (1998). The Berkeley FrameNet Project. Proceedings of COLING–ACL.
6. Jurafsky, D., & Martin, J. H. (2023). Speech and Language Processing (3rd ed., draft). Prentice Hall.

INTERNET SAYTLARI:

<https://umr-tool.cs.brandeis.edu/>

<https://n.ziyouz.com/portal-haqida/xarita/uzbek-nasri/ulug-bek-hamdami-1968>