

Moscow, June 14–17, 2023

## Representation of lexical polysemy in the database (semantic shift 'sun/day')<sup>1</sup>

Maxim Russo

Institute of Linguistics, Russian Academy of Sciences / 1 bld. 1 Bolshoy  
Kislovsky Lane, 125009 Moscow, Russia  
rousseau@iling-ran.ru

### Abstract

This paper, based on the data from more than 600 languages collected during the work on the database of semantic shifts in the languages of the world, addresses the semantic transition 'sun/'day'. We analyze the geographic and genealogical distribution of this semantic shift, the predominant direction of semantic development, and the patterns of morphological derivation associated with the shift.

**Keywords:** polysemy; lexical typology; database, semantic shift

**DOI:** 10.28995/2075-7182-2023-22-1108-1111

## Представление лексической полисемии в базе данных (семантический переход 'солнце'/'день')

Максим Руссо

Институт языкознания РАН / 125009, Москва, Большой Кисловский пер. 1, стр. 1  
rousseau@iling-ran.ru

### Аннотация

Работа, основанная на данных более 600 языков, собранных в процессе работы над базой данных семантических переходов в языках мира, посвящена семантическому переходу 'солнце'/'день'. Мы анализируем географическое и генеалогическое распространение данного семантического перехода, преобладающее направление семантического развития и связанные с переходом паттерны морфологической деривации.

**Ключевые слова:** полисемия; лексическая типология; база данных, семантический переход

### [1] Introduction

The presence of the meanings 'sun' and 'day' in same lexeme is one of the most common cases of polysemy in the languages of the world. For example Azerbaijani *günün tutulması* 'solar eclipse', *bir neçə gün qabaq* 'a few days ago'; Hungarian *delel a nap* 'the sun has risen', *minden nap* 'every day'; Chinese 日平西 / *rì píng xī* 'the sun tends to the west', 昨天是他值的日 / *zuótiān shì tā zhí de rì* 'yesterday was the day of his duty'.

However, probably due to the extreme rarity of this polysemy in the Indo-European languages, it did not attract active attention of researchers. The purpose of this work was to clarify the distribution of

---

<sup>1</sup> The study was supported by the Russian Science Foundation, project 22-18-00586 Mapping of the typology of polysemy via computer-assisted system of the cross-language identification of semantic shifts.

the polysemy 'sun' / 'day' in the languages of the world and to determine the semantic patterns associated with it, in particular, the direction of semantic development.

## [2] Material and Methods

The sample covered 652 languages representing 97 language families or isolated languages of Eurasia, Africa, Australia, Oceania, North and South America.

The work was carried out as part of the project Database of Semantic Shifts in languages of the world (DSS), Thus, it uses the understanding of the semantic shift as a cognitive proximity of two meanings that manifests itself in the fact that these two meanings are conflated within the limits of one word in an extended sense [15, 16]. In addition to the lexical polysemy, it also covers other types of realizations of the semantic shifts, in particular, morphological derivation and internal cognates. The realization of the transition in the form of external cognates occurs in cases where different meanings of a polysemantic proto-language word were lost in descendant languages (for example Atayal *wagi* 'sun' and Siraya *wagi* 'day' from Proto-Austronesian \*waRi<sub>1</sub> 'day; sun; dry in the sun'), and are not further considered in this article. The sample included cases where a semantic transition was found in one of the marginal designations of the sun or day in a language, although it was absent in the basic words (for example, Sanskrit *aruṣá* 'sun, day', initially 'red, reddish' in the presence of *sūrya* 'sun' and *dina* 'day').

All collected materials and sources of information used are available online at <https://clck.ru/33njWj>, data on the languages where the semantic shift was found are also included in the DSS (<https://datsemshift.ru/shift1007>).

## [3] Results

The semantic shift 'sun' - 'day' in the form of synchronic polysemy was found in 282 languages (43%, 100%=652), in 103 languages (16%) the shift exists as of derivation and internal cognates, 8 languages (0.01%) are obscure cases where such a shift may be suspected. In 248 languages (38%) the shift was not found.

The **polysemy** 'sun' - 'day' exists in the Indo-European (only Indo-Iranian, Greek, Tocharian), Uralic, Nakh-Dagestan (Dargin + Chamalal), Altaic (Turkic, Manchu, Korean, Japanese), Dravidian, Burushaski, Afro-Asiatic (excluding the Berber branch), Sino-Tibetan, Tai-Kadai, Miao-Yao, Austroasiatic (Bakhnar, Khasi, Khmer, Khmuic, Monic, Nicobaric, Pearic, Palaungic), Niger-Congo, Mande, Central Sudanese, East Sudanese, Komuz, Kulyak, Khoisan, Bilua, Engan, Kiwai, Manubaran, Maibrat, Sepik, Timor-Alor-Pantar, Torricelli, Trans-New Guinea, Yam, Pama-Nyunga, Muskogean, Mayan, Salish (only Lushucid), Mixe-Zoque, Tupian, Pano-Takanan, Caddoan, Uto-Aztec, Chonan, Chumashan, Macro-Jê, Chibchan, Panoan, Arawakan, Mascoian, Araucanian, Matacoan, Uru-Chipaya, Yuman-Cochimí, Hokan, Yana, Uti, and Candoshi-Shapra language. Only in the form of **derivation** or **cognates** the shift exists in Basque, Yeniseian, the Munda branch of the Austroasiatic languages, Saharan, Hadza, Malak-Malak, Na-Dene, Iroquois, Guahiboan, Zuni, Algic, Tukano, Siouan, Penutian, Jivaro, Purépecha, Pomoan, Nadahup, Huave, Maidu languages. The transition was not found in any form in the Abkhaz-Adyghe, Kartvelian, Chukchi-Kamchatkan, Nivkh, Yukagir, Ainu, Katuy, Mangic and Vietic branches of the Austroasiatic languages, Sandawe, Mairasi, Trans-Fly, Yale, Macro-Gunwinyguan, Tiwi, Southern Dayli, Eskimo-Aleut, Quechuan, Oto-Manguean, Nadahup, Aymaran, Haida, Tsimshianic, Zaparoan, Nambikwaran, Paez, Kutenai, Kunza. Geographically noteworthy is the relatively rare occurrence of the shift in the native languages of Canada and the United States, with the exception of the languages of the state of California. Also, the shift is rare in the northeast of Eurasia and in the west of Europe.

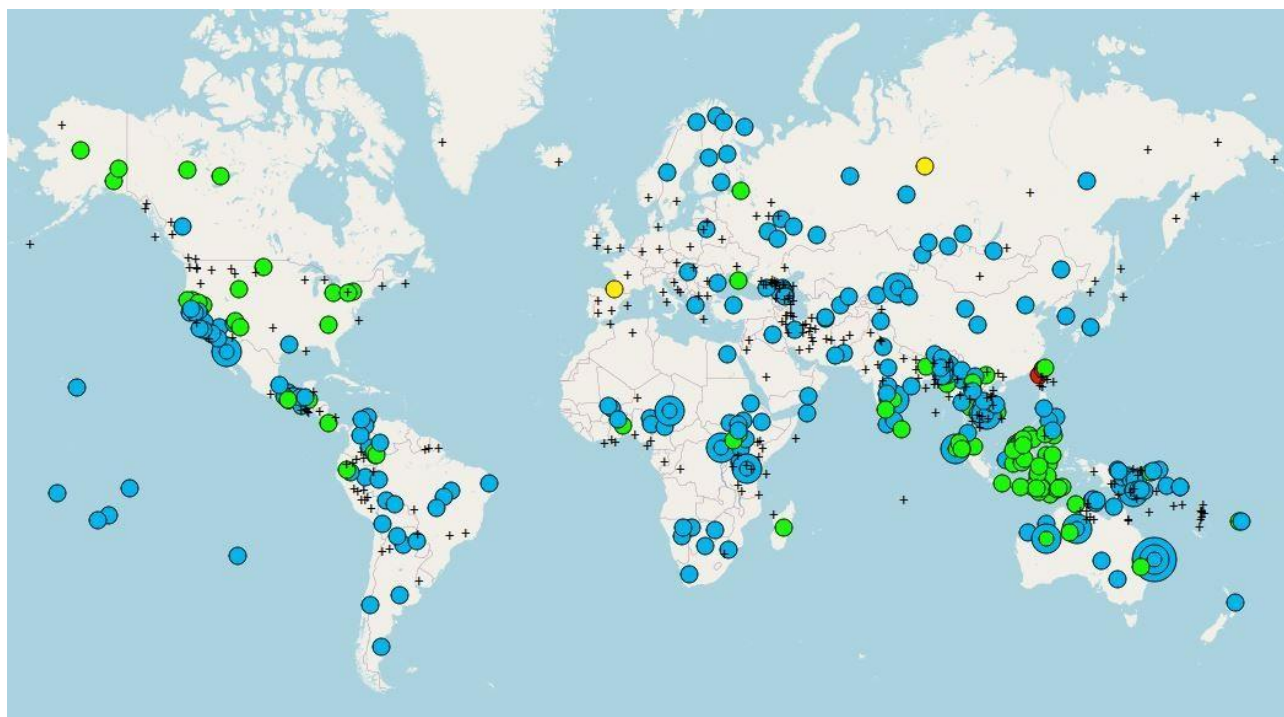


Figure 1: Languages with the semantic shift 'sun' — 'day' (blue – polysemy, green – derivation, yellow – internal cognates, red – external cognates, black cross – shift not found)

In some cases, the occurrence of this shift in individual languages in families, for which it is generally not typical, is explained by contact influence. As can be assumed, the Manchu word *šun* 'sun' (from Proto-Tungusic \**sigün* 'sun') acquired the meaning 'day' under the influence of the Chinese [3: II, 78]. This polysemy in Korean and Japanese can be explained by Chinese influence, although the for the Japanese cannot be completely excluded influence of the Austronesian substrate.

The shift is very rare in the Indo-European family, the noted examples in the Indo-Iranian languages and Ancient Greek are marginal designations for the day or the sun. In Ancient Greek, the word ἡλιος 'sun' in the meaning of 'day' is noted only in literary works and is extremely rare: ἡλίους δὲ μυρίους 'many days' (Euripides *Helene*, 652-653); ἄλιον ἄμφ' ἐνὶ 'during one day' (Pindar *O.* 13.37); Ἡλίουσ δέκα ἄσιτοι διαμενοῦμεν καὶ ἐπὶ παννύχουσ ὑμνωδίασ ἐπαγρυπνοῦντεσ ὄνειρώττομεν τὰ τοιαῦτα 'We, they said, do not take food for ten days, do not close our eyes, and after the all-night hymns see such dreams' (Pseudo-Lucianus *Philopatris*, 26).

In the basic lexemes of the Indo-European languages, the shift is observed in the Kurdish (Kurmanji), Baloch and Luristan of the Iranian group. In these languages, words with the polysemy 'sun' / 'day' go back to Proto-Iranian \**ráwčah* 'light', 'day', and are cognate to Avestian *raocah-*, Old Persian *raucah-* 'lamp', 'celestial body', 'light', 'daylight', 'day', Avestian *raok-* 'to shine', Middle Persian *rōz*, Classical Persian *rōz*, Farsi *ruz*, Tajik *рӯз*, Dari *ruz*, Parthian *rōž*, Zazaki *řōja*, *řōža*, *řōza*, Talysh *ruj*, *rūj*, *ruz*, *roz*, Tati *rōž*, Gurani *řō*, Gilaki *rūj*, Mazanderani *rōj*, *ruj*, Bakhtiari *res*, *ruz*, Vonishini, Kokhrudi, Zefre *rū*, Keshei *rū*, Sangisari *rūz*, Shamirzadi *rūz*, *ruz*, Sedei *rū*, Gazi, Kafroni *rū*, *rūz*, Sivandi *rū*, Semnani *rūz*, Sorkhei *rūz*, Lasgergi *rūz*, Pashto *wraž*, *wraz*, *rwaž* 'day [2: 217-218, 12: 104] It is also found in the Tocharian languages, where it can be considered a consequence of contacts with Chinese or Turkic languages.

The direction of semantic development 'sun' → 'day' occurs more often (15 out of 23 cases when it was possible to reasonably determine the direction of the transition), for example, in Soqotri *šam*, Akkadian *šamšu* 'sun, day' from Proto-Semitic \**šamš-* 'sun' [7: 418, 210, 10: 335-338, 6: 472]. It is noteworthy that in the two contact languages where this polysemy is noted, the meaning 'day' turns out to be an innovation, absent in the lexifier language: the Kyakhta Pidgin *соница* 'sun, day' from Russian *солнце* 'sun' [9: 337], Chinook Jargon *sun* [4]. However, there are languages where the meaning of 'sun' is later: Jibbali *yum* 'sun, sunlight, day' from Proto-Semitic \**yawm-* 'day' [5: 314] or Anatolian

Arabic *nahār* 'day' → 'sun' [1: 404]. It is common to use lexical modifiers to distinguish between sun and day. For example, in Seri *zaah* combined with the postpositive article *quij* 'compact', means 'the sun'; with the article *cop/cap* 'vertical; abstract' - 'the day'. This explains the cases when the words 'sun' and 'day' are internal cognates (Basque *eguzki* 'sun', *egun* 'day', Kubachi *бе* 'day', *бегвала* 'sun', Ket *u* 'sun', and *u'* 'day'), where differing formants can be fossilized morphemes.

In the case of derivation against the more frequent direction 'day' → 'sun' (93 out of 99 cases of the determined direction, see also [11: 291]). In India, both Indo-European and Dravidian languages, the sun's descriptive designations are "cause/lord/jewel/of the day" "day maker". A significant part of the cases of derivation (56 examples) is the model 'sun' = 'eye of the day', which is widespread in Southeast Asia [14]. Another frequent pattern of derivation (26 examples) is presented in the same word 'luminary' is used to designate both the sun and the moon, and if necessary, the expressions 'day luminary' and 'night luminary' are used, respectively. This colexification of the sun and moon is especially common in the Americas [13]. In some cases, the word 'luminary' may be lost, as in Cubeo *jāravu* 'day' → *aviá jāravucacu* 'sun' (lit. 'luminary of the day') → *jāravucacu* 'sun' (lit. 'of day') [8: 152, 355]. Rare examples of the direction 'sun' → 'day' are Tennet (South Sudan) *i*: 'sun', *i-tten* 'day', Malak-Malak (Australia) *miri* 'sun' and *mirien* 'day'.

## References

- [1] Behnstedt Peterand, Woidich Manfred Wortatlas der arabischen Dialekte. Band I: Mensch, Natur, Fauna un Flora. 2010. Leiden – Boston. Brill.
- [2] Cabolov Ruslan 2010. Kurdish etymological dictionary [Ètimologičeskij slovar' kurdsckogo jazyka]. Vol. 2. Moscow.
- [3] Cincius Vera (ed.) Comparative dictionary of the Tungusic languages. Materials for the etymological dictionary [Srvnitel'nyj slovar' tunguso-man'čžurskix jazykov. Materialy k ètimologičesckomu slovarju]. Leningrad, 1975
- [4] Gibbs, George, Dictionary of the Chinook Jargon, or Trade Language of Oregon. Washington: Smithsonian Institution, 1863.
- [5] Johnstone Thomas Muir. Jibbali Lexicon. Oxford - New York, 1981.
- [6] Kogan Leonid. Lexical evidence and the genealogical position of Ugaritic (I) // Babel und Bibel 3, 429-488
- [7] Leslau Wolf . Lexique Soqotri (sudarabique moderne) avec comparaisons et explications ètymologiques. Paris 1938.
- [8] Morse Nancy L., Salser Jay K., Salser Neva. Diccionario ilustrado bilingüe: cubeo-español español-cubeo. Bogota. Editorial Alberto Lleras Camargo, 1999.
- [9] Perexval'skaja Elena. Russian pidgins [Russkie pidžiny]. Saint Petersburg, 2008
- [10] Roth Martha T. (ed.).1992. Assyrian Dictionary of the Oriental Institute of the University of Chicago, Volume 17, S, Part 1.
- [11] Starostin Georgij. Languages of Africa: an attempt at a lexicostatistical classification. Vol. I. Methodology. Khoisan languages [Jazyki Afriki. Opyt postroenija leksikostatističesckoj klassifikacii. T. 1: Metodologija. Kojsanskie jazyki]. Moscow. 2013.
- [12] Schmitt Rüdiger. 2000 Die iranischen Sprachen in Geschichte und Gegenwart, Wiesbaden: Reichert Verlag
- [13] Urban Matthias. 2009. 'Sun' and 'Moon' in the Circum-Pacific Language Area // Anthropological Linguistics, Volume 51, Numbers 3-4, Fall and Winter 2009: 328-346.
- [14] Urban Matthias 2010. 'Sun' = 'Eye of the Day': A Linguistic Pattern of Southeast Asia and Oceania // Oceanic Linguistics, Volume 49, Number 2, December 2010: 568-579.
- [15] Zalizniak Anna A., Maria Bulakh, Dmitrij Ganenkov, Ilya Gruntov, Timur Maisak, and Maxim Russo. 2012. The catalogue of semantic shifts as a database for lexical semantic typology // Linguistics, 50(3):633–669.
- [16] Zalizniak Anna A., 2018. The Catalogue of Semantic Shifts: 20 Years Later // Russian Journal of Linguistics. 22, 4: 770-787.