

МЕТОДИ І АЛГОРИТМИ СУЧАСНИХ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

УДК 81-112.2; 81-112.4; 81'33

¹М. Білинський, ²О. Басалкевич

¹Львівський національний університет імені Івана Франка,

²Національний університет “Львівська політехніка”,
кафедра метрології, стандартизації та сертифікації

ДІАХРОННІ ДОСЛІДЖЕННЯ СИНОНІМІЇ З ВИКОРИСТАННЯМ РЕЛЯЦІЙНОЇ БАЗИ ДАНИХ

© Білинський М., Басалкевич О., 2011

Досліджено діахронні синонімічні ряди якісних прикметників на матеріалі давньошотландської мови. Зазначені ряди будуються та оцінюються за допомогою бази даних, створеної у середовищі Access. Прикладом слугує лексико-семантичне поле “sturdy-cowart”.

Ключові слова: метод антидатування, текстовий прототип, діахронний гіперонім, ваговий коефіцієнт, діахронна частота, діахронний синонімічний ряд якісних прикметників давньошотландської мови, біполярне діахронне лексико-семантичне поле.

The research of diachronic synonymic sets of qualitative adjectives of Older Scots is under consideration in the proposed article. The mentioned sets are formed and estimated with the help of the Access-configured database. The lexical semantic field “sturdy-cowart” is the issue of the suggested approach.

Key-words: antedating method, textual prototype, diachronic hypernym, weight coefficient, diachronic frequency, diachronic synonymic set of qualitative Older Scots adjectives, bipolar diachronic lexical semantic field.

Вступ. Теоретичне підґрунтя дослідження

У межах історичної лінгвістики досліджено синонімію у діахронному аспекті. Синонімічні ряди будуються на основі принципу діахронного текстового прототипу [3]. Методом антидатування, заперечення пізнішої дати більш ранньою встановлюється приблизний час появи слова. Дослідження проводиться на основі наявних друкованих джерел. Встановлена так лексична одиниця називається текстовим прототипом досліджуваного синоніма. Діахронний синонімічний ряд впорядковується у порядку зростання дати появи синоніма. Перший член ряду має назву діахронного гіпероніма. Вагомість кожного синоніма діахронного ряду визначається відповідним коефіцієнтом. Формула для його обчислення $W_i^{(y)} = \frac{y_i - y_{\min}}{y_{\max} - y_{\min}} \left(\frac{1}{n+1} - 1 \right) + 1$ [2] передбачає врахування діахронного діапазону ряду ($y_{\max} - y_{\min}$), тобто проміжку часу від появи найстаршого слова ряду (в даному випадку – гіпероніма) до наймолодшого, а також часової відстані до появи конкретного синоніма ($y_i - y_{\min}$) та кількості слів у синонімічному ряді n . Розподіл вагових коефіцієнтів діахронного синонімічного ряду нерівномірний, оскільки слова, що вперше вживаються у спільному джерелі, отримують однакові коефіцієнти.

Аналіз основних джерел

Подібне дослідження проводиться у [2] на матеріалі англійської мови. Назагал побудовані ряди слугують потенційними статтями історичного тезауруса. Відомо, що розроблення подібного

тезауруса англійської мови триває досі. Пропоноване у статті дослідження є певним вступом до аналогічних студій у давньошотландській мові, відгалуженні нортумбрського діалекту староанглійської мови, що самостійно розвивалася на теренах Шотландії у XII – XVI століттях. Матеріалом для дослідження є словник давньошотландської мови (Dictionary of the Older Scottish Tongue), який налічує 12 томів і разом зі словником сучасної шотландської мови (the Scottish National Dictionary) складає електронний словник шотландської мови (the Dictionary of the Scots Language), представлений у [5].

Принцип дослідження

Вихідним для кожного діахронного синонімічного ряду якісних прикметників давньошотландської мови у нашому дослідженні приймається синонімічний ряд англійських якісних прикметників, що вибирається з словників Oxford Learner's Thesaurus [4] чи WordNet [6]. Кожний член такого ряду розглядається як спільне значення для відповідних давньошотландських синонімів (третя ознака синонімії згідно з [1:6]). На основі поданих цитат перевіряється контекстуальна взаємозамінність синонімів (сьома ознака синонімії [1:6]). Встановлення принаймні двох ознак є достатньою умовою для формування синонімічного ряду. Щоб впорядкувати отриманий ряд, використовуємо описаний вище метод антидатування. Розглядаються подані у великій кількості ілюстративні цитати, на основі бібліографічних даних словника встановлюється цитата з найдавнішим датуванням, досліджувана лексична одиниця якої надалі вважатиметься текстовим прототипом синоніма, а дата – приблизним часом появи слова. На цьому етапі розроблена реляційна база даних ISET_intro виконує функції збереження і впорядкування отриманих даних (таблиці English_Strings, English_Quotations, Bibliography, Older_Scots_Synonyms, Older_Scots_Strings). На етапі формування діахронних синонімічних рядів, обчислення відповідних вагових коефіцієнтів, побудови діаграм діахронної частоти та виведення текстових прототипів використовуються запити I_Selection (формування необхідної вибірки), II_Selection_Year_Correction (корегування датувань бібліографії, якщо вказується конкретніша дата), III_Final_Selection (формування остаточної вибірки) та форми виводу Final_Selection_Parameters (для виводу параметрів ряду) та Final_Selection_Set (для виводу діахронного синонімічного ряду, розподілу вагових коефіцієнтів, років приблизної появи слів, текстових прототипів та означення джерел й авторства) цієї бази. Для роздруку вихідних даних розроблені відповідні звіти (English_set, OS_set, Diagram).

Побудова лексико-семантичного поля “sturdy-cowart”.

Результуючі дані для англійського синонімічного ряду: brave • courageous • heroic • gallant • fearless • gutsy (describe sb who is willing to do difficult or dangerous things, описують особу, охочу мати справу зі складними чи небезпечними речами) [4] подані у відповідній формі виводу, для унаочнення у конкретному випадку відображеній на двох рис. 1, 2.

У формі виведено впорядкований синонімічний ряд давньошотландських прикметників, значення яких відповідають синонімічному ряду англійських прикметників з гіперонімом “brave, хоробрий”, вказані роки появи кожного слова та відповідні вагові коефіцієнти, а також наведено цитати, що містять текстові прототипи вказаних синонімів з означенням джерела.

Отже, діахронний синонімічний ряд давньошотландських прикметників матиме вигляд: sturdy (1375: 1.00), kene (1375: 1.00), worthy (1375: 1.00), stalwart (1375: 1.00), wicht (1375: 1.00), joly (1375: 1.00), bald (1375: 1.00), derf (1375: 1.00), hardy (1375: 1.00), stout (1375: 1.00), muddy (1375: 1.00), manful (1400: 0.89), stern (1400: 0.89), vertuus (1400: 0.89), quent (1400: 0.89), hey-hartid (1420: 0.81), proud (1420: 0.81), wele hertit (1456: 0.65), hichty (1456: 0.65), coragious (1500: 0.46), lusty (1500: 0.46), unagast (1500: 0.46), gallant (1508: 0.43), martial (1513: 0.41), soverane (1513: 0.41), prati (1535: 0.31), heroick (1549: 0.25), magnanime (1549: 0.25), hartsum (1567: 0.17), valorous (1600: 0.03) [5]

Відповідні текстові прототипи:

He sa sturdy wes and stout That he wes the mast vorthy man That in-to Carrik liffit than – John Barbour, “The Bruce” (1375), With his men that war cant and kene [C. keyn] – Ibidem, He has in his cumpany Mony a gud man & worthi That sall be hard for till assay. – Ibidem, The dedys of stalwart folk that lywyt

ar – Ibidem, A *gheman wicht and hardy* – Ibidem, *Thai war all zong men and ioly*, And *zarnand till do cheuelry*. – Ibidem, *the king that stout wes, stark and bald* – Ibidem, *That wes derff, stout, and ek hardy* – Ibidem, A *gheman wicht and hardy* – Ibidem, *He wes bath zoung, stout and felloun, Richt ioly als and volageous*; – Ibidem, *He ... ourcom mony a mudy man* – Ibidem, *To loue men for virtue and nobilnes Syne honour thame for manfull excellence* – “The Buik of Alexander, or, The Buik of the Most Noble and Valiant Conqueror Alexander the Grit” (1400); *King Alexander the sterne and stout* – Ibidem, *Lo! heir thre wenchis full pietuous, And we ar thre, stark and vertuuous [F. vertuuous] ... The thre, I wait, weill luffis the tua [etc.]*; – Ibidem, *And zit, for all thare iolyte, The quentest [F. li plus cointe] sall abased be*; – Ibidem, *That wes a fulle hey-hartyd man* – Andrew of Wyntoun, “The Orygynale Cronykil of Scotland” (1420), *Downald Brec ... Kyng wes fourtene wyntyрrowde And eftyre that his dayis wes dwn*; – Ibidem, *That he be hardy and wele hertit to se justice done, and falde nocht his hert for myshertitnes of curage*; – Hay, G., “Gilbert of the Haye’s Prose Manuscript.” (1456); *Ane orguillous, hychty, and prydy rebelloure* – Ibidem, *Follow furth on Robyn Hude With hartis coragious & gud* – “Crying of Play, The Maner of the Crying of ane Play” (1500), *Behald our heidis, O lusty gallandis gay*; – The Poems of Robert Henryson (1500), *Full glaid thai glyde as gromes vnagaist*; – Craigie, W.A., “King Hart. In The Maitland Folio Manuscript” (1500); *Gif I be galland, lusty and blyth*; – The Poems of William Dunbar, (1508), *With curage marcyall* – Virgil’s Aeneid translated by Gavin Douglas. (1513), *Heir was ... The maist dowchty lynage sprang be kynde Fra Kyng Tewcer, campyones souerane [L. magnanimi heroes]*; – Ibidem, *Ane wonder prattie cheild* – William Stewart, *The Buik of the Croniclis of Scotland: or, a metrical Version of the History of Hector Boece. (1535), zour heroyque vertu*; – Lindsay, Sir David, *The Complaynte of Scotland. (1549), The immortal gloir that procedis ... fra zour magnanime auansing of the public veil*; – Ibidem, *Now euerie Douglas of ane hartsum mynde, Think on dame Margaret* – Sempill R, Davidson J., Maitland J. et al, *Satirical Poems of the Time of the Reformation. (1567) The valereus and godlie schiftane Judas Machabæus*; – Hamilton, John, *A Facile Traictise, contenant, first: ane infallible reul to discerne trew from fals Religion..etc. (1600) [5].*

OS_synonym	year	coeff	text prototype	title	author
kene	1375		1 With his men that war cant and ke	The Bruce.	John Barbour
worthy	1375		1 He has in his company Mony a gu	The Bruce.	John Barbour
stalwart	1375		1 The dedys of stalwart folk that lyw	The Bruce.	John Barbour
wicht	1375		1 A gheman wicht and hardy	The Bruce.	John Barbour
ioly	1375		1 Thai war all zong men and ioly, Ar	The Bruce.	John Barbour
mudy	1375		1 He ... ourcom mony a mody man	The Bruce.	John Barbour
derf	1375		1 That was derff, stout, and ek hard	The Bruce.	John Barbour
bald	1375		1 the king that stout wes, stark and	The Bruce.	John Barbour
sturdy	1375		1 He sa sturdy wes and stout That I	The Bruce.	John Barbour
stout	1375		1 He wes bath zoung, stout and fell	The Bruce.	John Barbour
hardy	1375		1 A gheman wicht and hardy	The Bruce.	John Barbour
quent	1400	0.89247311827957	And zit, for all thare iolyte, The qu	The Buik of Alexander, or, The Bu	
manful	1400	0.89247311827957	To loue men for virtue and nobilne	The Buik of Alexander, or, The Bu	
stern	1400	0.89247311827957	King Alexander the sterne and stc	The Buik of Alexander, or, The Bu	
vertuous	1400	0.89247311827957	Lo! heir thre wenchis full pietuous,	The Buik of Alexander, or, The Bu	
hey-hartid	1420	0.806451612903226	That was a fulle hey-hartyd man	The Orygynale Cronykil of Scotla	Wyntoun, Andrew of; Laing, Dav
proud	1420	0.806451612903226	Downald Brec ... Kyng wes fourte	The Orygynale Cronykil of Scotla	Wyntoun, Andrew of; Laing, Dav
wele hertit	1456	0.651612903225806	That he be hardy and wele hertit t	Gilbert of the Haye's Prose Manu	Hay, Gilbert of the; Stevenson, J
hychty	1456	0.651612903225806	Ane orguillous, hychty, and prydy	Gilbert of the Haye's Prose Manu	Hay, Gilbert of the; Stevenson, J
lusty	1500	0.462365591397849	Behald our heidis, O lusty galland	The Poems of Robert Henryson.	Henryson, Robert
unagast	1500	0.462365591397849	Full glaid thai glyde as gromes vn	King Hart. In The Maitland Folio M	Craigie, W.A.
coragious	1500	0.462365591397849	Follow furth on Robyn Hude With	Crying of Play, The Maner of the C	
gallant	1508	0.427956989247312	Gif I be galland, lusty and blyth;	The Poems of William Dunbar.	Dunbar, William
marcial	1513	0.406451612903226	With curage marcyall	Virgil's Aeneid translated by Gavr	Gavin Douglas
soverane	1513	0.406451612903226	Heir was ... The maist dowchty ly	Virgil's Aeneid translated by Gavr	Gavin Douglas
prati	1535	0.311827956989247	Ane wonder prattie cheild	The Buik of the Croniclis of Scott	William Stewart
heroick	1549	0.251612903225806	zour heroyque vertu	The Complaynte of Scotland.	Lindsay, Sir David
magnanime	1549	0.251612903225806	The immortal gloir that procedis ...	The Complaynte of Scotland.	Lindsay, Sir David
hartsum	1567	0.174193548387097	Now euerie Douglas of ane hartsum	Satirical Poems of the Time of the	Sempill, R, Davidson, J., Maitla
valorous	1600	0.032258064516129	The valereus and godlie schiftane	A Facile Traictise, contenant, fir	Hamilton, John

Рис. 1. Діахронний синонімічний ряд давньошотландських прикметників, відповідний синонімічному ряду англійських прикметників з гіперонімом “brave”

Наступний рисунок відображає діахронну частоту словотворення, тобто часовий розподіл кількості новостворених слів ряду.

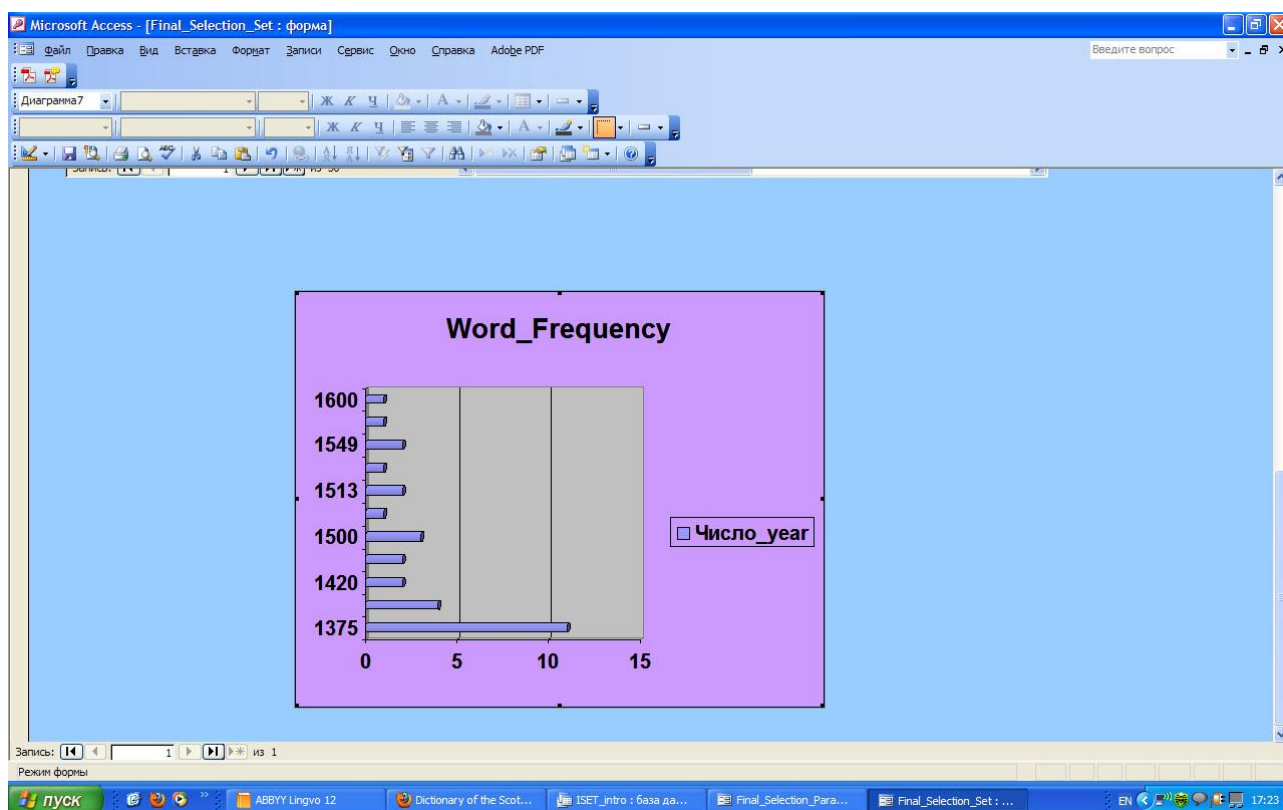


Рис. 2. Діахронна частота словотворення для синонімічного ряду давньошотландських прикметників, відповідного англійському синонімічному ряду *brave*

Очевидно, що пік словотворення припадає на 1375 рік, який датує появу роману Джона Барбора «Брюс». Діаграма відображає загальний стан речей: дослідження виявляють, що найбільша частка лексики давньошотландської мови належить до раннього періоду її формування на основі північного нортумбрського діалекту.

Антонімом слова “brave, відважний” є “cowardly, боягузливий” (пари антонімів вказуються у таблиці Antonyms), йому відповідає синонімічний ряд таких англійських прикметників:

cowardly • fearful (lacking courage; ignobly timid and faint-hearted, той, якому бракує мужності, боязкий і слабкодушій) [6] та діахронний синонімічний ряд давньошотландських прикметників (рис. 3):

cowart (1456: 1.00), spiritles (1500: 0.69), unhardy (1568: 0.21), undochty (1570; 0.2) [5]

*And he be **cowart** or full of wikkit vicis; – Hay, Gilbert of the; Stevenson, J.H., Gilbert of the Haye’s Prose Manuscript. (1456), The bernis both wes ... out of mesour marrit in thair mude As **spreitles** folkis on blonkis hvffit on hicht Both in ane studie starand still thai stude;- Craigie, W.A., King Hart. In The Maitland Folio Manuscript I lxxxvii (1500), It makis a perte mows Ane **vnhardy** catt;- Bannatyne, George, The Bannatyne Manuscript written in Tyme of Pest.(1568) *Quhat gart ʒow schute to slay ʒone men of gude? ... Deir sall ʒe by ʒone foule **vnduchtie** deid;* – Sempill R, Davidson J., Maitland J. et al, Satirical Poems of the Time of the Reformation. (1570) [5]*

Діаграма не виявляє жодної динаміки через обмеженість вибірки (рис. 3).

Синонімічні ряди антонімів формують лексико-семантичне біполярне поле. Отже, результатом продемонстрованого дослідження є історичний тезаурус діахронного лексико-семантичного біполярного поля “sturdy-cowart”.

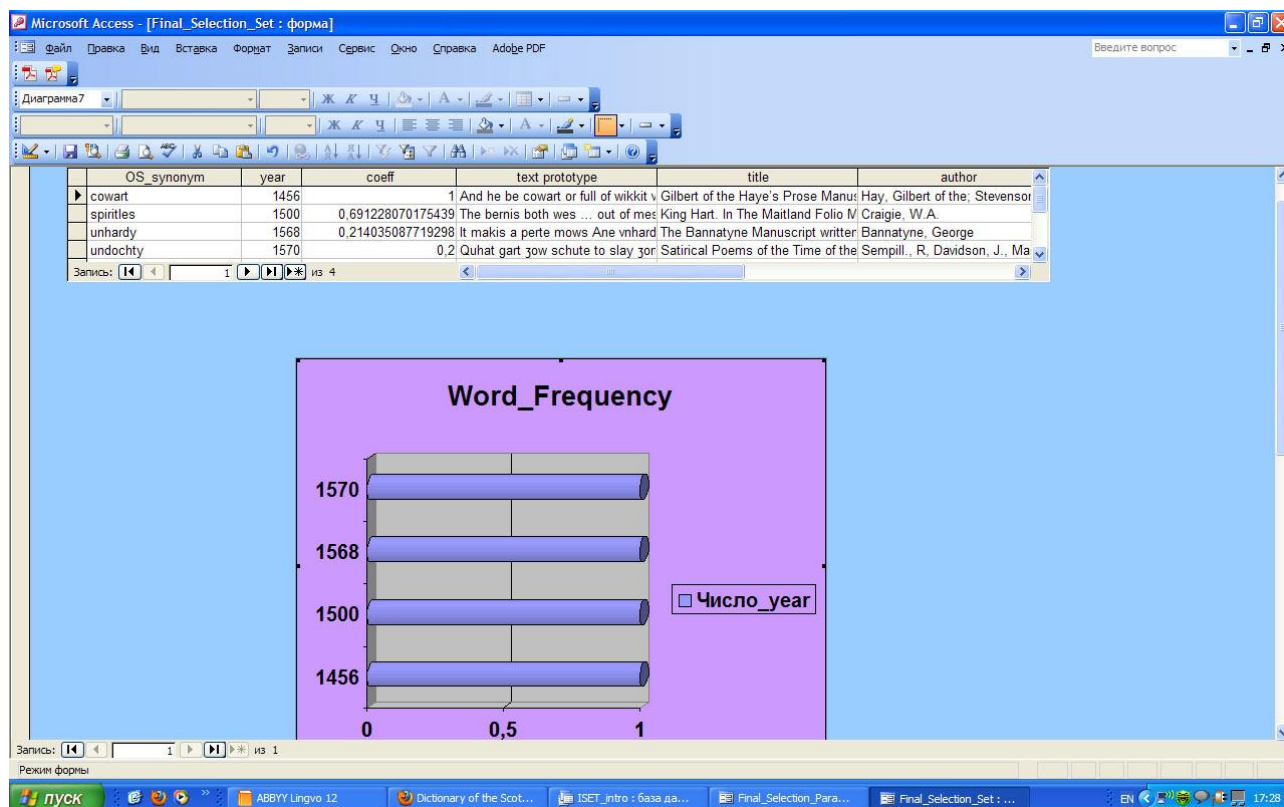


Рис. 3. Діахронний синонімічний ряд давньошотландських прикметників, відповідний синонімічному ряду англійських прикметників з гіперонімом "cowardly", та його частота словотворення

Висновки

Відображено алгоритм побудови діахронних синонімічних рядів давньошотландських прикметників, що використовується для створення бази даних часткового історичного тезауруса давньошотландських прикметників ISET_intro. Основу бази складають лексико-семантичні поля, одне з яких стало предметом викладеного у статті дослідження. Наступним кроком вбачається побудова зведених діаграм та дослідження кореляції між полями.

1. Вилюман В. Г. *Английская синонимика*. – М.: Высшая школа, 1980. — 128 с. 2. Altmann G., Bilynsky M., Pereumylida A. — *Weight factor formalisms in the study of lexical growth: the case of textually modeled strings of English verbs*. // *Методи аналізу тексту. Methods of Text Analysis*. Чернівці: ЧНУ, 2009. – С. 13–39. 3. Schafer, Jurgen. — *Documentation in the O.E.D.: Shakespeare and Nashe as test cases*. // *Oxford. Clarendon Press, 1980 — 176p*. 4. *Oxford Learner's Thesaurus, A dictionary of synonyms*. // *Oxford University Press., 2008 – 1008p*. 5. *The Dictionary of the Scots Language. Electronic source*. // *Edinburgh: Scottish Language Dictionaries, 2007. — www.dsl.ac.uk* 6. *WordNet. A lexical database for the English language. Electronic source*. // *Princeton University, 2006. — http://wordnet.princeton.edu*