

# 33 Visualising diachronic change in the collocational profiles of lexical near-synonyms

---

PRIMAHADI-WIJAYA-R., Gede & RAJEG, I Made  
Fakultas Sastra dan Budaya, Universitas Udayana

## Abstract

*Tracing the path of complex linguistic change in a dynamic and visual fashion could be seen as a daunting task. Nevertheless, the current presence of large diachronic corpora and cutting-edge computational visualisation tools have made it possible for linguists to concurrently visualise and observe linguistic change dynamically. In this contribution paper, a case study utilising one such tool called “motion charts” is presented. The main aim is to visualise diachronic change in the nominal collocational profiles of two synonymous temperature adjectives in English, i.e. hot and warm, over the last one and a half centuries (1860s-2000s) on the basis of data from the Corpus of Historical American English (COHA). Motion charts analysis on the two compared synonyms helps reveal, whether considered together or individually, several patterns of changes and stability of the distribution of the collocates.*

**Keywords:** *diachronic corpus linguistics, diachronic lexical semantics, diachronic corpora, COHA, motion charts, near-synonyms, collocation, temperature terms*

## 1 Introduction\*

The availability of robust diachronic corpora in recent years has provided linguists with a solid empirical ground to study changes of linguistic phenomena over time. In accordance with the development of such textual resources, research in diachronic corpus linguistics has been advancing towards exploring a range of statistical and visualisation techniques of considerable sophistication for analysing lexical and grammatical changes (Hilpert 2013a; Hilpert & Gries to appear). A recent paper by Hilpert (2011), for instance, has shown how a visualisation tool called *motion charts* (Gesmann & de Castillo 2014) can dynamically trace and present a holistic view of complex language change through time. The essence of linguistic motion charts is to visually represent quantitative linguistic data as a “series of diachronically ordered scatterplots” (Hilpert 2011, p.435). On the basis of the *Corpus of Historical American English* (COHA) (Davies 2012), Hilpert (2011) exemplified the method through two case studies on recent lexico-syntactic changes in American English. These case studies further demonstrate the substantial strength of motion chart analysis in handling

---

\* This paper is a contribution to a Festschrift in honour of Prof. Dr. I Gusti Made Sutjaja, commemorating at the same time the occasion of his 70<sup>th</sup> birthday and retirement on October 2<sup>nd</sup> 2014. Prof. Sutjaja’s total dedication to preserve Balinese language and culture, as prominently embodied in his extensive works, particularly on Balinese lexicography, has been truly inspirational for both of us as his former students. We would like to thank (in alphabetical order) Alice Gaby and Martin Hilpert for valuable feedback on the first draft of this paper. The usual disclaimers apply. Correspondence should be addressed to the first author at [primahadiwijaya@gmail.com](mailto:primahadiwijaya@gmail.com)

quantitative data of various complexities, be it two-dimensional or multi-dimensional data sets.

For the first case study, Hilpert (2011, pp.444–448) investigates diachronic change in the preferred syntactic categories among 119 frequent words displaying ambicategorical behaviours between nouns and verbs (e.g. *act/to act, love/to love*). The second case study further extends the use of two-dimensional data visualisation towards multi-dimensional data sets, in this case, six major syntactic sub-categorisation frames of English complement-taking predicates (Hilpert 2011, p.450). In order to reduce the complexity of such data sets so that they can be represented in a two-dimensional map in this case, a dimension-reduction technique called *Multi-dimensional Scaling* was applied (Hilpert 2011, p.448). To sum up, the two case studies indeed offer a unique and elegant methodological perspective to approach theoretical inquiries in historical linguistics. Be that as it may, alongside shifts in lexico-syntactic aspects, Hilpert (2011, p.458) also hints at other feasible areas of linguistic changes waiting for further studies to explore by means of the same visualisation technique.

The primary aim of this contribution is thus to take up one of Hilpert's (2011, p.458) practical pointers for extending motion chart analysis in language change to one of the central topics in lexical semantics, viz. lexical near-synonyms. Over the recent past, the study of lexical synonymy has benefited from the insights of usage-based linguistics, most notably corpus linguistics (see, e.g., Glynn 2010 for an overview). The key theoretical tenet in usage-based approach to lexical semantics as well as to meaning in general is that the “meaning of a word is reflected in the *linguistic* contexts in which it occurs frequently” (Stefanowitsch 2010, p.368, italics in original). This view entails an empirical line of research at our methodological disposal that uses naturalistic usage data—gathered from corpora for instance—as an empirical basis to reveal recurrent patterns reflecting semantic nuances of a word (cf. Glynn 2010; Stefanowitsch 2010; Tummers et al. 2005, *inter alia*). As one of the well-established research traditions in corpus linguistics, analysis of collocations, i.e. a word's lexical co-occurrence patterns, can be one particular way to operationalise the notion of “linguistic usage context” from which meanings of a word can be read off (Liu 2013; Stefanowitsch 2010). In the study of near-synonyms itself, measures of collocation have been amongst the most-frequently used techniques in differentiating meanings or usages associated with the synonymous elements (Gries & Stefanowitsch 2004; Stefanowitsch 2005; Liu 2013; Desagulier In press, *inter alia*).

In the same way, the present study investigates usages of two near-synonymous adjectives from the domain of TEMPERATURE, i.e. *hot* and *warm*, regarding diachronic change in their collocational profiles (cf. Hilpert 2008; Hilpert 2013b, pp.10–11). The focus is on the syntactic collocates of attributive *hot* and *warm*, namely the modified nouns. More specifically, the analysis is concerned with the distribution of different nominal collocates with respect to their co-occurrence ratios with *hot* and *warm*, and the way this pattern may or may not have changed across time.

Comparing the use of *hot* and *warm* in this way, hence, can disclose which nominal collocates are predominantly modified by *hot* or *warm* in the [Adj+N]

pattern, and which ones exhibit even proportions. Moreover, the inclusion of time windows in the analysis can ascertain whether *hot* and *warm* preserve an overall stability regarding their collocational preferences over time, including their ratios, or whether the two words experienced changes in their collocations. Another avenue of changes can be monitored through pinpointing the genesis of particular collocations, including their development thenceforth (e.g. whether or not they become increasingly frequent over time, and show increasing co-occurrence tendency towards one of the synonyms pair), as well as the disappearance of them. In short, it is possible to map the rise and fall of *hot* and *warm*'s respective collocational profiles with the technique applied herein.

Given that collocations may reflect the meanings of a word, and that the corpus this study refers to, i.e. COHA, reflects diachronic developments, it is simultaneously possible to discover patterns of semantic evolution of the two synonyms as evidenced by their collocational shift (Davies 2012, p.143). The inclusion of successive time frames in the study of collocational patterns of *hot* and *warm* in particular can add a diachronic contribution to recent corpus-based works on English temperature terms in general that instead used synchronic English corpora (e.g. Lorenzetti 2010; Rasulić 2010).

## 2 Data and methods

The data for this study were gathered from the *Corpus of Historical American English* (COHA) (Davies 2012), which is freely accessible on-line. COHA contains more than 400 million words and consists of texts from four genres: fiction, popular magazines, newspapers, and non-fiction books. It dates from 1810s to 2000s and is well balanced by genres in each decade (Davies 2012, p.122f.). Additionally, COHA is carefully lemmatised and tagged for Part-of-Speech (POS). This particular feature facilitates a more restrictive query on certain words based on their part-of-speech. Since newspapers entered COHA from 1860s onwards, following Hilpert (2011, p.438), the data prior to 1860s were not included.

This study made use of the “compare” and “POS tag” features provided by COHA to specifically retrieve the lemmatised form of the collocating noun directly following *hot* and *warm*—their R1 collocate—framed in the [Adj+N] pattern, as in *hot water* and *warm welcome*. The minimum cut-off frequency of each collocate type was set to five<sup>1</sup>, and collocates included in this study are only those making their way into the top one-hundred collocates. After setting up and submitting the search queries, COHA then returned the results in the form of a comparison table showing how often a particular noun collocates with *hot* and/or *warm*. Searches with the same parameters described previously were repeated for each decade, starting from the 1860s to the 2000s. The search outputs were then imported into spreadsheet software to be post-processed for input data (see Table 1) required for generating motion chart analysis in

---

<sup>1</sup> This value is the default set by COHA and this study simply uses that default setting. However, for the sake of the user, that setting is easily modified.

the form of scatterplots (see Figure 1 & Figure 2)<sup>2</sup>. Since the corpus in every COHA period are of unequal sizes, this study normalised the raw frequency of co-occurrence of collocates into frequency per million words so that meaningful comparison of frequency across the periods can be made (Gries 2010, p.271).

**Table 1** Snippet of the data frame input for a motion chart

Decade	Noun collocate	Collocate frequency with <i>hot</i> (x-axis values)	Collocate frequency with <i>warm</i> (y-axis values)	Combined frequency
1860	<i>admirer</i>	0.00	0.41	0.41
1860	<i>affection</i>	0.00	0.93	0.93
1860	<i>afternoon</i>	0.23	0.29	0.53
1860	<i>air</i>	0.64	1.05	1.69
...	...	...	...	...

Table 1 above captures only a small subset of the data from the 1860s. Further down, each row of the complete table holds all data points for the following decades. The x- and y-axis values correspond to the co-occurrence frequencies per million words of each collocate with *hot* and *warm* respectively, and provide the coordinate points for all of the collocates on the plot. A combined frequency value provides an input for computing the circle radius used to represent the overall area of each bubble in Figure 1 and Figure 2, while the sizes of the bubbles have been scaled to be proportional with the plotting area (Kabacoff 2011, p.278f.).

### 3 Results and discussion

Figure 1 shows the visual output of transforming the quantitative information presented in Table 1 into the so-called bubble plot, which is essentially a special version of a scatterplot (Kabacoff 2011, p.278f.). The graph in Figure 1 visualises co-occurring nominal collocates of *hot* and *warm* in the COHA decade of the 1860s. The x-axis designates the co-occurrence frequency per million words of collocates with *hot* whereas the y-axis portrays their co-occurrence frequency per million words with *warm*. Each bubble symbolises the nominal collocate whose coordinate position is determined from its respective co-occurrence frequency with *hot* and *warm*. Moreover, the different sizes of the bubbles, which are derived from the combined frequencies, indicate that the larger the bubble, the higher its co-occurrence usage ratio with *hot* or *warm*, or even with both. When looking at the bubble sizes in Figure 1, it is apparently obvious that *blood*, *heart*, *water*, and *weather* are the most frequent modified noun collocates. There are forty-nine nominal collocate types altogether in the 1860s but only twenty are labelled for readability reason. The labelled collocates are rendered into the dark grey bubbles while the unlabelled ones translate into the light grey bubbles.

<sup>2</sup> All computations and graphical analyses in this paper have been performed with the open-source programming language and statistical software R (R Core Team 2014). The data files and R program scripts to reproduce the graphics in this paper are available from the first author upon request.

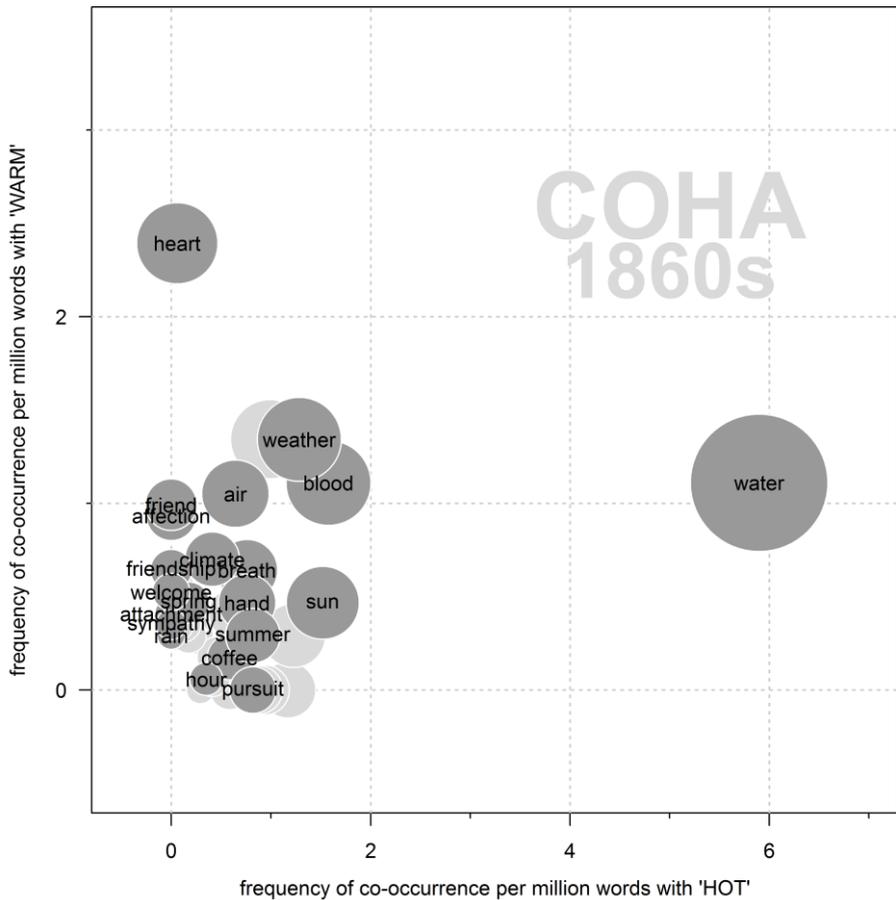


Figure 1 Co-occurrence frequencies of noun collocates of *hot* and *warm* in the 1860s

In summary, Figure 1 suggests that *hot* and *warm*, to a certain extent, exhibit clear preferences for the types of nominal collocates they typically modify, albeit the two synonyms roughly equally denote ‘an escalated temperature state’. Nouns such as *affection*, *attachment*, *friend*, *friendship*, *heart*, *rain*, and *welcome* clearly display a close affinity towards *warm*, meanwhile *hour*, *pursuit*, *summer*, *sun*, and *water*, including the unlabelled *haste* and *blast*, exemplify predominant collocates of *hot* in the 1860s. Nevertheless, there are also a number of shared collocates quite evenly modified by *hot* and *warm*, for instance *blood*, *breath*, *hand*, and *weather*. Interestingly enough, though, the majority of the respective predominant nominal collocates of *warm* especially, and of *hot*—as in the nominal phrases *hot pursuit* and *hot haste*—instantiate their metaphorical usages: those collocating with *warm* mostly denote ‘friendliness’ sense, while those with *hot* refer to ‘excitement’ (e.g. *hot haste*) and ‘intense’ (*hot pursuit*) senses. ‘Friendliness’ and ‘excitement’ senses have also been identified as amongst the most frequent ones characterising the metaphorical usages of the respective words in a recent study by Lorenzetti (2010, pp.6–8) on the basis of

synchronic English corpora, i.e. the *British National Corpus* (BNC) and the *Corpus of Contemporary American English* (COCA).

At this point, the question of whether *hot* and *warm* have undergone change in their collocational profiles from the 1860s to the 2000s can only possibly be addressed by reproducing the graph from Figure 1 into a series of scatterplots (cf. Figure 2) for each subsequent decade starting from the 1860s onwards. Identical information as in Figure 1 is mirrored in the x- and y-axes of the graphs in Figure 2. There are total 265 nominal collocate types<sup>3</sup> over the span of one and a half centuries, but only nine items are highlighted and hence labelled: *bath*, *day*, *dog*, *heart*, *pursuit*, *smile*, *spot*, *water*, and *welcome*. They have been selected since they provide cases of variation and stability to be discussed in the remainder of this paper.

To begin with, it appears immediately evident that the lower left of the graphs in Figure 2 has been the most populated area in which most collocates cluster throughout the periods, hence reflecting the relatively low co-occurrence frequency per million words of the collocates. Nevertheless, after examining the selected elements, it is found that they have gone through some observable development.

First, while *heart* is mainly the preferred collocate of *warm* in the 1860s and lasts for the rest of the decades, its usage level with *warm* has gone down steadily since then. This includes its disappearance from the graphs between 1970s and the 1990s<sup>4</sup> before having eventually reappeared by the end of the fifteen decades. Moreover, similar declining trend can be observed in *welcome*. After a short period of fluctuation towards reaching its heyday in the 1890s, *welcome's* usage ratio with *warm* has gradually dropped since the 1900s<sup>5</sup> but it has kept its definite preference towards *warm* stable over time. Furthermore, since getting into the picture in the 1930s, *smile* has also collocated almost exclusively with *warm* in an upward trend.

<sup>3</sup> The following items are all nominal collocates of *hot* and/or *warm* under analysis: *admirer, affection, afternoon, air, alkali, anger, appreciation, area, argument, arm, ash, attachment, baby, bath, bed, beer, bird, biscuit, blanket, blast, blood, body, box, bread, breakfast, breath, breeze, brick, broth, brow, bun, butter, button, cake, car, chase, check, cheek, chick, chicken, chili, chocolate, cider, clasp, climate, cloak, cloth, clothes, clothing, coal, coat, cocoa, coffee, color, commodity, corn, corner, country, cross, cup, darkness, date, day, debate, desert, discussion, dish, dog, drink, drop, earth, embrace, evening, eye, face, fat, feel, feeling, fight, filament, finger, fire, fish, flame, flash, flesh, flush, food, foot, forehead, friend, friendship, fudge, fur, gas, girl, glow, glue, good, grasp, grease, griddle, grill, ground, guy, hand, haste, head, heart, hour, house, hug, indignation, interest, imitation, iron, issue, item, jazz, kiss, kitchen, lava, lead, lick, light, line, lip, liquid, lunch, mama, market, meal, metal, milk, money, month, morning, music, mustard, nest, news, night, number, oil, one, oven, palm, pan, pant, part, pavement, pepper, period, pie, pixel, place, plate, platter, pot, potato, praise, pressure, property, pudding, pursuit, rain, ray, reception, regard, right, rinse, robe, rock, rod, rodder, roll, room, rum, sand, sauce, sausage, scent, scotch, sea, season, seat, seller, set, sex, shawl, sheet, shit, shot, shower, side, skillet, skin, sleep, smell, smile, smoke, solution, soup, sox, spell, spot, spring, springs, steam, stock, stone, stove, streak, stream, stuff, sud, suit, sulphur, summer, sun, sunlight, sunshine, supper, support, supporter, surface, sympathy, syrup, tamale, tar, tea, tear, temper, thing, thrill, ticket, time, tin, tip, toddy, tone, topic, touch, towel, tribute, tub, valley, voice, waffle, war, wash, water, wave, wax, weather, welcome, wheel, wind, wine, winter, wire, woman, word, work, zone*

<sup>4</sup> Spot check in COHA decade of the 1970s to the 1990s has revealed that the disappearance of *heart* from the graph is due to its co-occurrence frequencies that are below the minimum cut-off frequency of five set in this study. The followings are co-occurrence frequencies (in bracket) of *heart* with *hot* and *warm* in the three decades mentioned: *warm heart* (4) vs. *hot heart* (0), 1970; *warm heart* (3) vs. *hot heart* (0), 1980; *warm heart* (3) vs. *hot heart* (0), 1990. Despite its absence, it is clear that *heart* still is the common collocate of *warm* compared to *hot* in the respective decades.

<sup>5</sup> Similarly to *heart*, the absence of *welcome* from the graph in the 1940s also owes to its surprisingly low co-occurrence frequency with *warm* in the [Adj+N] pattern (only two tokens), which, in fact, is the overall lowest one across the periods of 1830 to 2000.

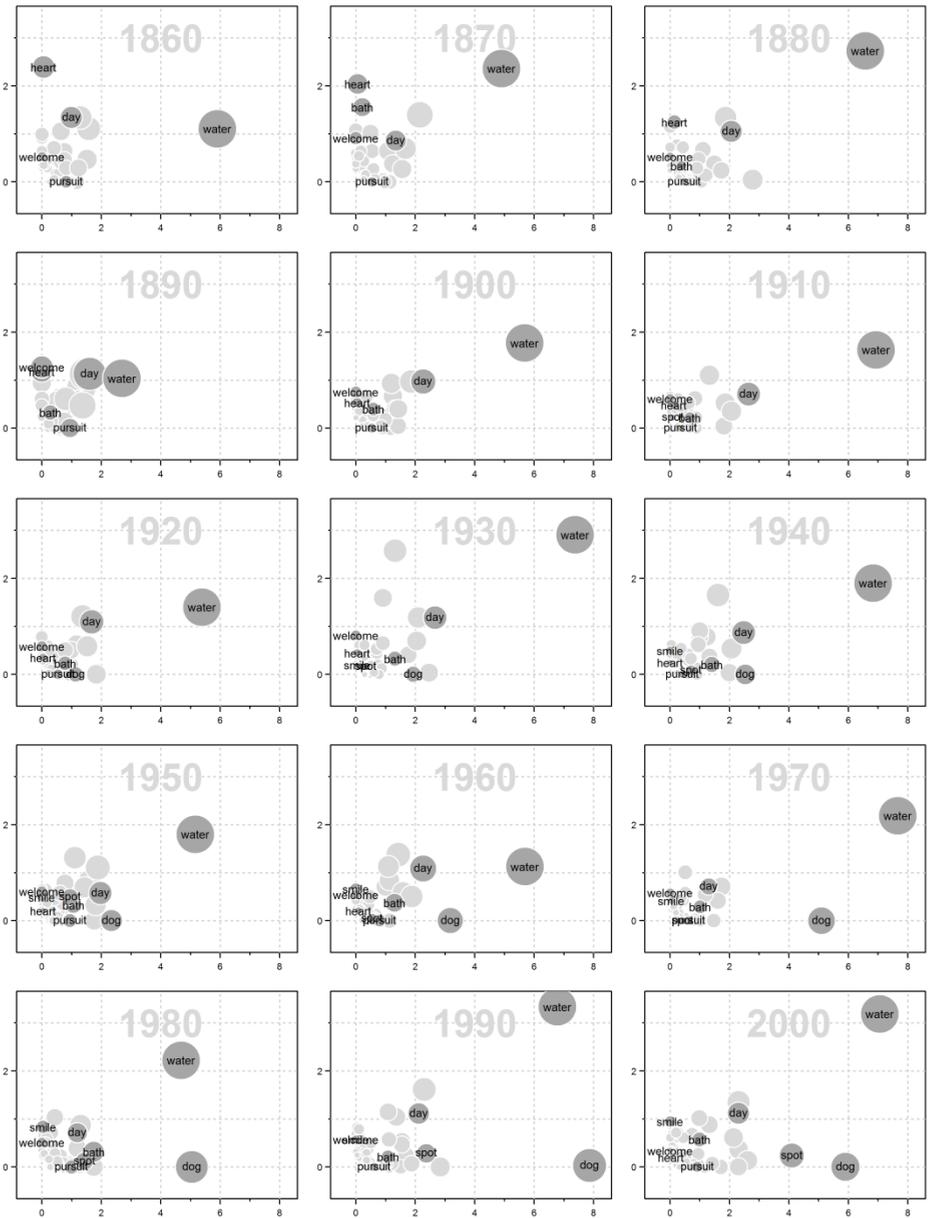


Figure 2 Changes in the collocational profiles of *hot* and *warm*, COHA 1860s-2000s

The decades-long association displayed by the three abovementioned collocates with *warm*, which accentuates *warm*'s metaphorical sense of 'friendliness', offers a historical dynamics on Lorenzetti's finding (2010, p.8) regarding the synchronic status of this metaphorical sense as the most frequent one associated with *warm* in English.

On the reverse side, though, several observations regarding the development of the collocational profiles of *hot* can also be made. The item *water*, for instance, being the most frequent and preferred nominal collocate of *hot* overall, has shown a roughly positive trend in its usage ratio with both *hot* and *warm* as the time goes by. This trend

is especially more noticeable towards *warm* rather than *hot*. Furthermore, *spot* has also exhibited an escalated preference over *hot* in a doubling-up fashion by the 1980s onwards, after between 1940 and 1970 co-occurring only somewhat more often with the same word rather than with its synonym. The noun *dog* is another case in point. In contrast with *spot* to some extent, *dog* has built its exclusive collocational profile with *hot* right from the start in the 1920s, and the ratio of this collocational pattern has risen gradually, not even slowing down towards its peak in the 1990s, followed by nothing but a modest slip-back in the 2000s. The noun *pursuit* meanwhile indicates an instance of consistency in its marked collocation with *hot*.

The inclusion of *hotdog* and *hotspot* herein—considered more as compounds nowadays rather than regular [Adj+N] noun phrases—, in particular, illustrates how diachronic frequency leap may induce lexicalisation of certain syntagmatic phrases through univerbation. Univerbation, as a subtype of lexicalisation, is the process by which a syntactic phrase or construction coalesces into a single (compound) word (Brinton & Traugott 2005, p.47f.). The argument over the potential role of frequency strength for triggering such process via diachronic language change can be attributed to Bybee's, saying that the strength of sequential relations developed by two or more words "is determined by the frequency with which the two words appear together" (Bybee 2010, p.25). In addition, Haas (2012, p.18) even more explicitly states that lexicalisation of a collocation into a univerbated expression arises from its frequent co-occurrence (cf. Schmid 2007, p.121). This implies that lexicalisation via univerbation involves an increase in relative dependency of the component structures of a complex expression to yield a compressed structured-whole of form-and-meaning pairing with a relatively high degree of unit-hood. Moreover, as a result of tacitly frequent repetition, those univerbated sequences of words would be fostered to become established or, in another technical parlance, institutionalised multi-word expressions (cf. Brinton & Traugott 2005, pp.45–47; Schmid 2007, p.121; Langacker 2008, p.21f.; Bybee 2010, p.35).

Institutionalisation of a particular expression encompasses both (i) the acceptance of the expression by part or all of certain speech community, and (ii) its limitation to denote specific meaning orientation, which would not usually be strictly predictable—but, rather, be motivated—from the component parts of the complex expression (Brinton & Traugott 2005, p.45f.; cf. Langacker 2008, p.169f.). Then, the established or institutionalised expression is also part of social institution, in the sense of serving as "a conventional label for a conventional concept" (Brinton & Traugott 2005, p.47). The expressions such as *hotdog* and *hotspot* well exemplify these phenomena. Firstly, their specialised meaning orientations as a unit denoting 'a sausage' and 'a place with wireless internet connection' respectively have turned the meaning of the adjectival component *hot* less salient, even though it has still contributed indirectly (e.g. exploited via metaphorical or metonymical mechanisms [cf. Benczes 2006]) to the expressions' composite meanings. Secondly, they have also been used as label of conventional concepts in recent times, i.e. a well-known food as part of

cultural change in American society, and a widespread technology-related concept of wireless internet connection.

Another type of change that can be identified from Figure 2 is an inter-synonym shift in usage proportion of particular collocates. This is well exemplified by *bath* and *day*. At the beginning of the period, *bath* collocates almost exclusively and rather frequently with *warm* prior to a sharp fall in the 1880s. Thereby, *bath* has moderately drifted towards *hot* and has constantly maintained its lower usage ratio with *warm* despite its minor increase in the 2000s. By the 2000s, hence, *bath* prefers to co-occur with *hot*, albeit just slightly. By the same token, *day* initially co-occurs a little more often with *warm*, but, for the rest of the periods, it has proceeded fluctuatively to be rather more frequently used with *hot*. While space does not allow further in-depth discussion on the cause of collocational change in *hot* and *warm* and the way it reveals about their semantics, yet brief general comment on causes of language change follows.

From the usage-based perspective, the cause of language change is not a matter of language *per se*, but more essentially a matter of the people's actions when they use the language (Evans & Green 2006, p.123). This view was developed long ago by a German linguist, Hermann Paul, in his *Prinzipien der Sprachgeschichte* (1920 [1880], p.32) saying "The real cause of the change of (linguistic) conventions is nothing other than ordinary language use." (quoted from Haspelmath 1999, p.1066, endnote 17). Moreover, according to William Croft's *Utterance Selection Theory* (see Evans & Green 2006, p.123), there are two-step processes through which people change the way language is used, i.e. "innovation" and "propagation" (cf. Traugott & Trousdale 2013, p.91f.). The former could include the need for expression coinage to catch up with, for instance, technological advances (Evans & Green 2006, p.124) (*hotspot* is a case in point), or simply just the need for being socially successfully noticed by being "extravagant" with the innovative use of particular expressions. This is what is termed as "Maxim of Extravagance" (see Haspelmath 1999 for more details). The later process, i.e. propagation, involves the spread and habituation of the new usage across the community through frequent use before becoming fully conventionalised (Evans & Green 2006, p.123f.). For instance, the increasingly frequent use of collocational pattern of *hot* with nouns such as *girl*, *woman*, *chick*, or *guy* in the 2000s<sup>6</sup>, conveying the '(sensually) attractive' sense of *hot*, might exemplify the way that particular sense is being fostered to become more conventionally associated with *hot* in recent years. In addition, according to Haspelmath (1999, p.1058), the frequent use of particular lexical items can be due to their usefulness in fulfilling a discourse function. This could also be the case for those collocations of *hot* including their '(sensually) attractive' sense.

After all, a sequential series of static scatterplots shown in Figure 2, without a doubt, can serve to reveal quite a great deal of information regarding linguistic change, which otherwise would be difficult to convey. However, this paper is also accompanied with online material<sup>7</sup> in which readers can find the animated version of the graphs in

<sup>6</sup> *hot girl* (1940 [1 token]; 1950 [1]; 1990 [1]; 2000 [7]); *hot woman* (1990 [2]; 2000 [5]); *hot chick* (1990 [1]; 2000 [11]); *hot guy* (1980 [2]; 1990 [2]; 2000 [18]).

<sup>7</sup> URL to the online motion chart resources: <http://primahadiwijaya.blogspot.com/2014/07/motion-charts-resources.html>

Figure 2, in addition to other useful resources. The advantages of such version of the graph as compared to the static ones, among other things, are twofold. First, readers can interactively inspect changes taking place in specific elements that interest the readers, and in the remaining unlabelled elements herein. Second, and more interestingly, the chart can be played out and watched as a movie showing dynamic process of diachronic language change.

#### 4 Conclusion

Building on Hilpert's study (2011), this paper has further explored the potential of *motion charts analysis* into another suggested field of linguistic inquiry, viz. lexical near-synonyms. In particular, this study explores diachronic change in the nominal collocational profiles of loosely synonymous pair of temperature adjectives, i.e. *hot* and *warm*. The application of *motion charts analysis* in this study has allowed effective identification of different patterns of changes in particular collocates with respect to their usage distributions with *hot* and *warm*. The noticeable changes include changes in usage ratio of particular collocates of a word as well as inter-synonyms shift in collocational preferences. Changes in the collocational profiles can inform not only the historical study of the semantics of *hot* and *warm* in this case (e.g. *hot pursuit*, *warm smile*), but they can also shed light on the cultural association of certain collocations (e.g. *hotdog*).

#### References

- Benczes, R., 2006. *Creative Compounding in English. The Semantics of Metaphorical and Metonymical Noun-Noun Combination*, Amsterdam & Philadelphia: John Benjamins.
- Brinton, L.J. & Traugott, E.C., 2005. *Lexicalization and Language Change*, Cambridge: Cambridge University Press.
- Bybee, J., 2010. *Language, Usage and Cognition*, Cambridge: Cambridge University Press.
- Davies, M., 2012. Expanding horizons in historical linguistics with the 400-million word Corpus of Historical American English. *Corpora*, 7(2), pp.121–157.
- Desagulier, G., In press. Visualizing distances in a set of near synonyms: *rather*, *quite*, *fairly*, and *pretty*. In D. Glynn & J. Robinson, eds. *Corpus Methods for Semantics. Quantitative studies in polysemy and synonymy*. Amsterdam: John Benjamins, pp. 145–178.
- Evans, V. & Green, M., 2006. *Cognitive Linguistics: An Introduction*, Edinburgh: Edinburgh University Press.
- Gesmann, M. & de Castillo, D., 2014. *Using the Google Chart Tools with R: googleVis-0.5.2 Package Vignette*, Available at: <http://cran.r-project.org/web/packages/googleVis/vignettes/googleVis.pdf> [Accessed June 16, 2014].
- Glynn, D., 2010. Corpus-driven Cognitive Semantics. Introduction to the field. In Dylan Glynn & Kerstin Fischer, eds. *Quantitative Methods in Cognitive Semantics: Corpus-Driven Approaches*. Berlin: Mouton de Gruyter, pp. 1–41.
- Gries, S.T., 2010. Useful statistics for corpus linguistics. In Aquilino Sánchez & Moisés Almela, eds. *A mosaic of corpus linguistics: selected approaches*. Frankfurt am

- Main: Peter Lang, pp. 269–291. Available at:  
[http://www.linguistics.ucsb.edu/faculty/stgries/research/2010\\_STG\\_UsefulStats4CorpLing\\_MosaicCorpLing.pdf](http://www.linguistics.ucsb.edu/faculty/stgries/research/2010_STG_UsefulStats4CorpLing_MosaicCorpLing.pdf) [Accessed July 18, 2014].
- Gries, S.T. & Stefanowitsch, A., 2004. Extending collocation analysis: A corpus-based perspective on “alternations.” *International Journal of Corpus Linguistics*, 9(1), pp.97–129.
- Haas, F., 2012. Motivating an English-German contrast in word-formation. Available at:  
[http://www.personal.uni-jena.de/~fo45nar/files/einander\\_paper.pdf](http://www.personal.uni-jena.de/~fo45nar/files/einander_paper.pdf) [Accessed July 31, 2014].
- Haspelmath, M., 1999. Why is grammaticalization irreversible? *Linguistics*, 37(6), pp.1043–1068. Available at:  
[https://www.academia.edu/3015784/Why\\_is\\_grammaticalization\\_irreversible](https://www.academia.edu/3015784/Why_is_grammaticalization_irreversible) [Accessed September 04, 2014].
- Hilpert, M., 2013a. *Constructional Change in English: Developments in Allomorphy, Word Formation, and Syntax*, Cambridge: Cambridge University Press.
- Hilpert, M., 2013b. Corpus-based Approaches to Constructional Change. In T. Hoffmann & G. Trousdale, eds. *The Oxford Handbook of Construction Grammar*. Oxford Handbook Online. Oxford: Oxford University Press, pp. 1–13. DOI: 10.1093/oxfordhb/9780195396683.013.0025
- Hilpert, M., 2011. Dynamic visualizations of language change. Motion charts on the basis of bivariate and multivariate data from diachronic corpora. *International Journal of Corpus Linguistics*, 16(4), pp.435–461.
- Hilpert, M., 2008. *Germanic Future Constructions: A Usage-based Approach to Language Change*, Amsterdam: John Benjamins.
- Hilpert, M. & Gries, S.T., to appear. Quantitative approaches to diachronic corpus linguistics. In M. Kytö & P. Pahta, eds. *The Cambridge Handbook of English Historical Linguistics*. Cambridge: Cambridge University Press. Available at:  
[http://www.linguistics.ucsb.edu/faculty/stgries/research/ToApp\\_MH\\_STG\\_QuantMethInDiachrCorpLing\\_CUPHb.pdf](http://www.linguistics.ucsb.edu/faculty/stgries/research/ToApp_MH_STG_QuantMethInDiachrCorpLing_CUPHb.pdf) [Accessed June 14, 2014].
- Kabacoff, R.I., 2011. *R in Action. Data analysis and graphics with R*, Shelter Island, New York: Manning.
- Langacker, R.W., 2008. *Cognitive Grammar: A Basic Introduction*, Oxford: Oxford University Press.
- Liu, D., 2013. Salience and construal in the use of synonymy: A study of two sets of near-synonymous nouns. *Cognitive Linguistics*, 24(1), pp.67–113.
- Lorenzetti, M.I., 2010. Temperature Terms and Their Metaphorical Usages in English and Italian: A Contrastive Study. In *Linguistic Agency LAUD*. General & Theoretical Papers. LAUD 2010. University of Duisburg, Essen: LAUD, p. A763. Available at:  
[http://www.linse.uni-due.de/laud-downloadliste/articles/that-girl-is-hot-her-dress-is-so-cool-and-im-just-chilling-out-now-variation-in-metaphorical-usages-of-temperature-terms-in-engl.html?file=tl\\_files/laud/A763.pdf](http://www.linse.uni-due.de/laud-downloadliste/articles/that-girl-is-hot-her-dress-is-so-cool-and-im-just-chilling-out-now-variation-in-metaphorical-usages-of-temperature-terms-in-engl.html?file=tl_files/laud/A763.pdf). [Accessed June 29, 2014].
- Rasulić, K., 2010. What’s hot and what’s not in English and Serbian: A contrastive view on the semantic extensions of temperature adjectives. Available at:

[http://temperature.ling.su.se/images/a/a4/Rasulic\\_Serbian.pdf](http://temperature.ling.su.se/images/a/a4/Rasulic_Serbian.pdf) [Accessed November 12, 2013].

- R Core Team, 2014. *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria. Available at: <http://www.R-project.org/> [Accessed April 10, 2014].
- Schmid, H.-J., 2007. Entrenchment, Saliency, and Basic Levels. In D. Geeraerts & H. Cuyckens, eds. *The Oxford Handbook of Cognitive Linguistics*. Oxford: Oxford University Press, pp. 117–138.
- Stefanowitsch, A., 2010. Empirical cognitive semantics: Some thoughts. In Dylan Glynn & Kerstin Fischer, eds. *Quantitative Methods in Cognitive Semantics: Corpus-Driven Approaches*. Berlin: Mouton de Gruyter, pp. 355–380.
- Stefanowitsch, A., 2005. The function of metaphor: Developing a corpus-based perspective. *International Journal of Corpus Linguistics*, 10(2), pp.161–198.
- Traugott, E.C. & Trousdale, G., 2013. *Constructionalization and Constructional Changes*, Oxford: Oxford University Press.
- Tummers, J., Heylen, K. & Geeraerts, D., 2005. Usage-based approaches in Cognitive Linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory*, 1(2), pp.225–261.