Jarosław Wiliński

A covarying-collexeme analysis of the German caused-motion construction in the soccer domain

Acta Philologica nr 46, 17-24

2015

Jarosław Wiliński

A covarying-collexeme analysis of the German caused-motion construction in the soccer domain

1. Introduction

In recent years, there has been a growing interest in the investigation of the relationship between words and grammatical structures. Some studies have focused on the preferences or restrictions associated with individual slots in the construction (Stefanowitsch and Gries, 2003; Gries and Stefanowitsch, 2004a). Others examined potential interactions between (sets of) words occurring in two different slots of the same construction: for example, the into-causative, English possessive constructions, and the way-construction (Gries and Stefanowitsch, 2004b; Stefanowitsch and Gries, 2005). So far, however, little attention has been paid to the investigation of correlations between lexical items occurring in two different slots of the German caused-motion construction. The existing research has concentrated on the investigation of the association between the caused-motion construction in English and the words occurring in a particular slot provided by it. This paper seeks to investigate correlations between lexical items occurring in two different slots of the German caused-motion construction in the domain of soccer, and to indicate that such correlations are determined by frame-semantic knowledge. The reminder of this paper is organized as follows. In section 2, theoretical and methodological issues are addressed. Section 3 discusses data, material, and tools. In section 4, a statistical procedure is described. Section 5 provides an account of the caused-motion construction in the soccer domain. The results are presented in section 6. Section 7 gives an evaluation of the results and makes some proposals for future research.

2. Theoretical and methodological frameworks

This study adopts the terminology and the basic assumptions of Construction Grammar, an approach to grammar which holds that both lexical units and grammatical structures are meaningful. This theory, associated with the works of Goldberg (1995, 2003, 2006) and Croft (2001, 2007), views the construction as the basic unit of linguistic organization. The term *construction* is defined as a symbolic unit (Langacker, 1987), a pairing of form and meaning, where both form and meaning are construed broadly: the former includes morphemes, lexical units, syntactic structures, and even phonological and prosodic units, while the latter covers semantics, discourse function, and also social and cultural parameters of use.

A corpus-based method, referred to as *covarying collexeme analysis* (cf. Gries and Stefanowitsch, 2004b, Stefanowitsch and Gries, 2005), is used to investigate correlations between lexical items occurring in two different slots of the same construction.

In contrast to traditional colligation or collocation-based methods, it is applied to determine the frequency of a lexeme in one slot of a given construction in relation to the lexemes occurring in a different slot of the same construction. In other words, this approach aims to identify which words are strongly attracted or repelled by particular slots in a construction (that is, occur more frequently or less frequently than expected). Words that are attracted to a particular construction are referred to as *collexemes* of this construction (Stefanowitsch and Gries, 2003, pp. 214–215).

3. Data, material, and tools

The major source of data used in this paper comes from specialized corpora, covering the years between 2005 and 2013. Corpora include different types of texts (such as official news, comments, biographies, written interviews with people linked with the sport, match reports and reviews, etc.) derived from various internet websites. The German corpus includes ca. 1 000 700 words of contemporary written German soccer news. A software program, MonoConc Pro, was used to retrieve the observed frequencies from the corpus. The rest of the figures and expected frequencies were worked out by means of Microsoft Excel spreadsheets. All values required for the calculation of the association strengths were entered in the 2-by-2 table and submitted to the Fisher exact test. The p-value provided by this test was regarded as an indicator of association strength, i.e., a word's strength of attraction/repulsion to a construction: the smaller the p-value, the stronger the association. This statistical analysis was carried out by means of on-line Fisher's exact test calculator for two-by-two contingency tables.

4. Procedure, data retrieval and evaluation

Let us illustrate this procedure using the verb *setzen* in the verb slot together with the noun *Kopfball* in the noun slot of the direct object in the caused-motion construction. The first step of this procedure involves working out the observed frequencies. These are calculated in the following way: first, all occurrences of the caused-motion construction are identified from the corpus: 556. Second, the frequency of the verb *setzen* in the verb slot is determined: 44. Third, the frequency of the noun *Kopfball* in the direct-object slot is extracted: 27. Finally, the frequency of the verb *setzen* and the object *Kopfball* is counted: 17. These four values are derived from the corpus directly while the remaining figures (the frequency of all other verbs and object *Kopfball*: 10, the frequency of verb *setzen* and other object nouns in the caused-motion construction: 27, the frequency of all other verbs and other object nouns in the caused-motion construction: 502, the total frequency of all other object nouns in the caused-motion construction: 529, the total frequency of all other verbs in the caused-motion: 512) result from addition and subtraction.

The expected frequency of the verb *setzen* and the object *Kopfball* in the caused-motion construction is calculated as follows. The total frequency of the verb *setzen* in the caused-motion construction (44) is multiplied by the total frequency of the object *Kopfball* (27). Then this figure is divided by the total frequency of the caused-motion construction (556), giving the result (2.14). If the observed frequency of verb *setzen* and

its direct object *Kopfball* in the caused-motion construction is significantly higher or lower than expected, the association between the verb *setzen* and the direct object *Kopfball* is one of attraction or repulsion respectively (the verb *setzen* and the direct object *Kopfball* are then perceived as significantly attracted or repelled collexemes of the caused-motion construction).

Another step of this procedure involves working out the association strength of the verb *setzen* and the direct object *Kopfball*. In order to calculate the association strength between these pairs of lexical items, the following four frequencies need to be employed: the observed frequency of the verb *setzen* and the direct object *Kopfball* in the caused-motion construction; the frequency of all other verbs and the direct object *Kopfball* in this construction; the observed frequency of the verb *setzen* and other direct objects in the construction; the frequency of all other verbs and other direct objects in the construction. These are entered in a two-by-two table and examined by means of the Fisher exact test. The p-value resulting from the calculation of Fisher exact test for this distribution is exceptionally small: 3.49E-14. Comparing the observed frequency of the verb *setzen* and its direct object *Kopfball* with the expected one, we can notice that this verb occurs more frequently than expected with the direct object in the construction. In other words, *setzen* and *Kopfball* are highly significant, very strongly attracted collexemes of the caused-motion construction. Such results become meaningful only when this procedure is applied to every relevant pair of words in the caused-motion construction. In the next step of the procedure, the pairs are sorted according to direction of association and strength of association. Finally, the data are interpreted in a variety of ways. Suffice here to say that there are indeed pairs that are significantly attracted or repelled to each other in caused-motion construction, and that relations between two different slots of the same construction are determined by frame-semantic knowledge and image schemas.

5. Caused-motion construction

Let us now consider the caused-motion construction, which has been analyzed in detail by Goldberg (1995). This construction can be represented structurally as [SUBJ [V OBJ OBL]], where V is a dynamic verb and OBL is a directional phrase. OBL, which stands for "oblique," is realized in the caused motion construction as a directional prepositional phrase, as can be evident from the examples (a-c) below:

a) *Magnien spielte eine lange Flanke von links auf Hilbert* (Magnien played a long pass from the left side to Hilbert).
b) *Ribéry spielte den Ball direkt in die Mitte* (Ribéry played the ball directly into the middle).
c) *Zwei Minuten später setzte Ronaldo einen Kopfball an die Latte* (Two minutes later Ronaldo steered a header onto the crossbar).

The basic meaning of this construction is that a causer or agent directly causes a theme to move along a path designated by the directional phrase: namely, X CAUSES Y to MOVE Z.

The various extensions from this basic sense can be found in the domain of soccer. For example, in addition to the schematic CAUSED-MOTION frame, the afore-mentioned

examples are structured by two more specific frames: either the SHOT frame or the PASS frame. In the former case, a shooter causes a ball to move in a particular direction from a source location on the pitch along a path to a target location by hitting it in a certain manner with a body part. In the latter case, a passer causes a ball to move in a certain direction from a source location on the pitch along a path to a goal location, or a teammate. The subject in both cases denotes a player. The direct object refers to a ball or an action (e.g. *Ball*, *Freisto*β, *Flanke*, *Kopfball*, etc.). The directional complement is a prepositional phrase that is realized by the following frame elements: Source, Path, Goal, Recipient.

These semantic descriptions enable us to predict roughly what verbs and nouns are expected to occur in both slots of this construction. The verbal slot should prefer verbs denoting the acts of moving a ball or shooting at a target. The object slot should prefer nouns denoting a ball or an action. These predictions for a covarying-collexeme analysis of two different slots of the caused-motion construction in the soccer domain will be tested below. According to the principle of semantic compatibility (Stefanowitsch and Gries, 2005), lexical items can (or are likely to) occur in two different slots of the same construction if (or to the degree that) their meaning is compatible with the semantics provided by the construction for those slots. Clearly, this principle does not explain the kind of semantic coherence we can expect for those slots. In this study, it will be argued that the semantic coherence between two different slots of the caused-motion construction in the soccer domain is determined by frame-semantic knowledge. In other words, if two words refer to a situation in which a shooter causes a (moving) ball to move in a certain direction from the source location on the pitch along a path to a target (a goal location) by hitting it with a body part, it is likely that the semantic coherence is based on the SHOT frame. If in turn they describe a situation in which a passer causes a ball to move in a certain direction from the source location on the pitch along a path to a goal location, or a teammate, then we could expect the semantic coherence based on the PASS frame.

6. Results and discussion

The results confirm the predictions made about the semantic coherence between two different slots of the same construction and the specific suggestions concerning the meaning of both specific extensions of the caused-motion construction. Consider now Table 1, which shows the twenty most strongly attracted collexeme pairs for the caused-motion construction in German. The data in this table reveals an interesting tendency. The collexeme pairs appear to be ordered so that the top of the list features combinations being specific instances of the two major sub-senses of the construction: namely, to cause a ball to move in a certain direction from the source location on the pitch along a path to a target (a goal location) by hitting it in a particular manner with a body part or to cause a ball to move in a certain direction from the source location on the pitch along a path to a goal location, or a teammate.

	a	x	e	z	b	c	y	f	d	(a)	$P_{Fisher\ exact}$
Kopfball setzen	17	27	44	556	10	27	529	512	502	2.14	3.49E-14
Volley abziehen	4	5	4	556	1	0	551	552	551	0.04	1.27E-09
Pass spielen	5	7	17	556	2	12	549	539	537	0.21	2.88E-07
Freistoβ treten	6	47	7	556	41	1	509	549	508	0.59	1.76184E-06
Freistoβ zirkeln	6	47	7	556	41	1	509	549	508	0.59	1.76184E-06
Warnschuss abgeben	3	4	6	556	1	3	552	550	549	0.04	2.79636E-06
Hereingabe schicken	3	13	3	556	10	0	543	553	543	0.07	1.00378E-05
Schuss lenken	9	35	31	556	26	22	521	525	499	1.95	4.09851E-05
Abpraller stochern	3	10	5	556	7	2	546	551	544	0.09	4.13206E-05
Flanke schlagen	4	21	7	556	17	3	535	549	532	0.26	4.93428E-05
Ball schieben	20	241	25	556	221	5	315	531	310	10.84	0.000151001
Schuss abwehren	4	35	6	556	31	2	521	550	519	0.38	0.000181879
Ball bringen	30	241	44	556	211	14	315	512	301	19.07	0.000482981
Ecke bringen	4	7	44	556	3	40	549	512	509	0.55	0.001008556
Lupfer parieren	2	3	13	556	1	11	553	543	542	0.07	0.001496549
Elfer verwandeln	1	1	1	556	0	0	555	555	555	0.00	0.001798561
Fehlpass zurückspielen	1	1	1	556	0	0	555	555	555	0.00	0.001798561
Eckball wuchten	2	7	6	556	5	4	549	550	545	0.08	0.001992872
Freistoβ boxen	3	47	4	556	44	1	509	552	508	0.34	0.002140567
Maβflanke köpfen	2	2	27	556	0	25	554	529	529	0.10	0.002274937

a=Observed Frequency of verb (e.g. *setzen*) and object (e.g. *Kopfball*) in caused-motion construction; **b** = Frequency of all other verbs and object (*Kopfball*) in caused-motion construction; **c** = Observed Frequency of verb (*setzen*) and other object nouns in caused-motion construction; **d** = Frequency of all other verbs and other object nouns in caused-motion construction; **e** = Total frequency of verb (*setzen*) in caused-motion; **f** = Total frequency of all other verbs in caused-motion; **x** = Total frequency of object (*Kopfball*) in caused-motion construction; **y** = Total frequency of all other object nouns in caused-motion construction; **z** = Total frequency of caused-motion construction; **(a)** = Expected frequency of verb (*setzen*) and object (*Kopfball*) in caused-motion construction; $\mathbf{P}_{\mathbf{Fisher\ exact}}$ = index of co- varying collostructional strength

Table 1. The results of co-varying collexeme analysis for caused-motion construction

It can be seen from Table 1 that the former sense is instantiated by the collexeme pairs such as *Kopfball setzen, Volley abziehen, Warnschuss abgeben, Abpraller stochern, Flanke schlagen, Ball schieben* that evoke the SHOT frame. The Fisher-Yates Exact p-values for these combinations are 3.49E-14; 1.27E-09; 2.79636E-06; 4.13206E-05; 4.93428E-05; 0.000151001, indicating that the associations between two words are relatively strong ones. Note that *Kopfball setzen* is the most strongly associated covarying-collexeme pair in this construction. The latter sense in turn is represented by concrete instances such as *Pass spielen, Hereingabe schicken* which refer to the PASS frame. Their p-values resulting from the calculation of Fisher-Yates Exact are 2.88E-07 and 1.00378E-05 respectively. Fourth-ranked *Freisto*β *treten* and fifth-ranked *Freisto*β *zirkeln*, which belong to the DEAD BALL POSITION frame, can also be instances

of either the 'PASS' frame or the SHOT frame. In addition to co-varying collexeme pairs mentioned above, there are also combinations (e.g. *Schuss lenken, Schuss abwehren, Lupfer parieren*) that evoke either the PLAYER'S INTERVENTION frame or the GOALKEEPER'S INTERVENTION frame (in which an intervening player or a goalkeeper causes a moving ball to move into a particular locative goal by changing its direction) and pairs of words (such as *Ball bringen* and *Ecke bringen*) indicating the TO DELIVER PASS frame (in which a player succeeds in transferring a ball into a locative goal or to a teammate). There is also one pair (*Elfer verwandeln*) that appears to instantiate the relation between the TO TAKE OPPORTUNITY frame and the SHOT frame (i.e. a player takes an opportunity to score a goal by shooting the ball into the target location).

As follows from Table 2, the five less significantly attracted collexeme pairs for the caused-motion construction in the German language are *Ball setzen, Ball lenken, Freisto*β *setzen, Ball kratzen, Ball schlagen*, since their p-values resulting from the calculation of Fisher exact are very high: 0.997370418; 0.988374043; 0.982618275; 0.981804514; 0.981804514 respectively. Additionally, comparing the observed and the expected frequencies for each pair of words confirms that these collexeme pairs occur less frequently than expected in the caused-motion construction. Thus, they are strongly repelled pairs of covarying (Verb + Direct Object) collexemes in this construction.

	a	**x**	**e**	**z**	**b**	**c**	**Y**	**f**	**d**	**(a)**	$\mathbf{P}_{\text{Fisher exact}}$
Ball verlängern	6	241	15	556	235	9	315	541	306	6.50	0.69793606
Freistoβ droschen	1	47	14	556	46	13	509	542	496	1.18	0.714026934
Ball abfälschen	2	241	5	556	239	3	315	551	312	2.17	0.719394237
Leder bringen	2	37	44	556	35	42	519	512	477	2.93	0.812153794
Leder setzen	2	37	44	556	35	42	519	512	477	2.93	0.812153794
Ball schaufeln	1	241	3	556	240	2	315	553	313	1.30	0.818904775
Ball holen	1	241	3	556	240	2	315	553	313	1.30	0.818904775
Kugel setzen	2	38	44	556	36	42	518	512	476	3.01	0.82419617
Flanke setzen	1	21	44	556	20	43	535	512	492	1.66	0.828750264
Ball ablegen	1	241	4	556	240	3	315	552	312	1.73	0.897826925
Freistoβ bringen	2	47	44	556	45	42	509	512	467	3.72	0.905482206
Ball hämmern	3	241	11	556	238	8	315	545	307	4.77	0.921766965
Ball stochern	1	241	5	556	240	4	315	551	311	2.17	0.942435097
Schuss setzen	1	35	44	556	34	43	521	512	478	2.77	0.949291662
Ball abwehren	1	241	6	556	240	5	315	550	310	2.60	0.967613213
Ball schlagen	1	241	7	556	240	6	315	549	309	3.03	0,981804514
Ball kratzen	1	241	7	556	240	6	315	549	309	3.03	0.981804514
Freistoβ setzen	1	47	44	556	46	43	509	512	466	3.72	0.982618275
Ball lenken	8	241	31	556	233	23	315	525	292	13.44	0.988374043
Ball setzen	11	241	44	556	230	33	315	512	282	19.07	0.997370418

Table 2. The twenty most strongly repelled collexeme pairs of caused-motion construction in German

7. Conclusion

In this paper the analysis of the co-variation of collexemes in different slots of the German caused-motion construction has been carried out. In order to perform this analysis, the perspective of construction grammar and the corpus-based method, referred to as covarying collexeme analysis, have been adopted.

The results of this study indicate that in the case of the German caused-motion construction in the soccer domain, the semantic coherence between the covarying collexemes rests on frame-semantic knowledge. That is, if two words evoke a situation in which a shooter causes a (moving) ball to move in a certain direction from the source location on the pitch along a path to a target (a goal location) by hitting it with a body part, then we can find the semantic coherence based on the SHOT frame. If in turn they refer to a situation in which a passer causes a ball to move in a certain direction from the source location on the pitch along a path to a goal location, or a teammate, then we can find the semantic coherence based on the PASS frame. There are also some examples of covarying collexeme pairs whose semantic coherence is based on the PLAYER'S INTERVENTION frame, the GOALKEEPER'S INTERVENTION frame, the DEAD BALL POSITION frame and TO DELIVER PASS frame.

The method employed in this paper may be used in applied linguistics and language pedagogy for many specific purposes. First, since this approach yields more accurate results than traditional collocate-based-methods, it may increase the precision of language description. Second, the covarying collexeme analysis can be used to identify the most significant pairs of words from the perspective of their learning and teaching by determining which lexical items or pairs of lexical items are strongly associated with or repelled by a particular construction. Finally, it can be employed for developing linguistic theory. The practical applications of this method are much wider, and future research into different constructions will indicate its full potential.

References

Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective.* Oxford: Oxford University Press.

Croft, W. (2007). Construction grammar. D. Geeraerts and H. Cuyckens (Eds.), *The Oxford handbook of cognitive linguistics* (463–508). Oxford: Oxford University Press.

Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure.* Chicago: Chicago University Press.

Goldberg, A. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Science*, *7*(5), 219–224.

Goldberg, A. (2006). *Constructions at work: The nature of generalization in language.* Oxford: Oxford University Press.

Gries, S. Th. and A. Stefanowitsch. (2004a). Extending collostructional analysis: A corpus-based perspective on alternations. *International Journal of Corpus Linguistics*, *9*(1), 97–129.

Gries, S. Th. and A. Stefanowitsch. (2004b). Co-varying collexemes in the into causative. M. Achard and S. Kemmer (Eds.), *Language, culture, and mind* (225–236). Stanford, CA: CSLI.

Langacker, R. (1987). *Foundations of cognitive grammar* (Vol. 1). Stanford, CA: Stanford University Press.

Stefanowitsch, A. and S. Th. Gries. (2003). Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics*, *8*(2), 209–243.

Stefanowitsch, A. and S. Th. Gries. (2005). Covarying collexemes. *Corpus Linguistics and Linguistic Theory*, *1*(1), 1–43.

Streszczenie

Celem niniejszego artykułu jest zaprezentowanie wyników badań ilościowych i jakościowych nad konstrukcją wyrażającą ruch spowodowany przyczyną zewnętrzną w domenie piłki nożnej. Wyniki pomiarów istotności statystycznej leksemów występujących w dwóch różnych miejscach tej samej konstrukcji pokazują, że istnieją pary leksemów, które są mocno i słabo powiązane z tą konstrukcją. Ponadto uzyskane wyniki potwierdzają tezę, że semantyczna spójność pomiędzy najbardziej znaczącymi parami leksemów jest determinowana przez ramowo-semantyczną wiedzę.

Abstract

This paper adopts a constructional approach to grammar and the collostructional method, which is designed to investigate interactions between lexical items occurring in two different slots of the same grammatical structure. The method, referred to as "covarying collexeme analysis," is applied to identify the association strength between pairs of words occurring in two different slots of the German caused-motion construction in the soccer domain. The results of this investigation reveal that the relations between two different slots of this construction are determined by frame-semantic knowledge.