

---

## Coding asymmetries, frequency and predictability: the case of *to* vs *from*

---

Mittwoch  
08.03.2017  
14:45 - 15:15  
B4 1, 0.26

Laura Becker  
*Universität Leipzig*

[laura.becker@uni-leipzig.de](mailto:laura.becker@uni-leipzig.de)

Matías Guzmán Naranjo  
*Universität Leipzig*

[matias.guzman\\_naranjo@uni-leipzig.de](mailto:matias.guzman_naranjo@uni-leipzig.de)

Different asymmetries with respect to the frequency/complexity/length of an expression from the domain of spatial relations have been claimed to hold cross-linguistically. One such case is the goal-source asymmetry (eg. Nam 2004, Stefanowitsch and Rohde 2004, Lakusta 2005, Lakusta and Landau 2005, Markovskaya 2006, Regier and Zheng 2007, Papafragou 2010): the goal being more relevant and prominent, it has been argued that motion to is expressed more frequently and in a less complex way than motion from. Different motivating mechanisms, e.g. iconicity, markedness, or diachronic development, have been proposed to be responsible for the attested effect. An alternative that has recently gained popularity for explaining different kinds of coding asymmetries is that these asymmetries are the product of frequency effects (Haspelmath 2008). In a coding asymmetry, the less frequent form will be at least as complex (in length or number of markers, etc.) as the more frequent form. The explanation given is that the more frequent forms are more predictable, and thus they need to be less marked for speakers to identify them.

However, this explanation requires two hypotheses that have not yet been explored. First, it assumes that higher frequency always translates into higher predictability, and second, it assumes that more predictable forms will be shorter. So far, typological work on coding asymmetries has not directly explored these claims.

Addressing the goal-source asymmetry, we propose an empirical approach to the relation between frequency, iconicity, and predictability. We show that frequency and predictability of an expression are not linked in a trivial way, since the latter depends on frequencies of both the whole expression (verb + preposition) and the individual frequencies of verbs and prepositions. Results also show that the semantics of the verbal context influence the predictability of the motion type, so that iconicity could play a role independent from frequency.

In order to address the expression of motion to and from in Spanish and Portuguese, we selected 12 verbs that can combine with prepositions expressing

both motion types. The verbs were divided into two groups: (i) those that feature inherent deictic semantics and imply directionality (Spanish *ir* ‘go’, *venir* ‘come’, *llegar* ‘arrive’, *salir* ‘exit, leave’, *entrar* ‘enter’, *llevar* ‘take’, *traer* ‘bring’, as well as their Portuguese counterparts); (ii) manner motion verbs that are neutral to directionality and deixis (Spanish *andar* ‘stroll’, *correr* ‘run’, *caminar* ‘walk’, *cargar* ‘carry’, *viajar* ‘travel’, *nadar* ‘swim’, as well as the Portuguese counterparts). As for the expression of direction, we included the prepositions *a* and *para* for motion to, as well as *de* and *desde* coding motion from.

We extracted the frequency for each verb + preposition pair from the Corpus del Español Web and Corpus do Português Web (Davies 2015–2016, Davies and Ferreira 2015–2016), and for the individual frequencies of the verb and preposition on their own. We then calculated the directional attraction, known as  $\Delta p$ , from the verb to the preposition (Gries 2013). This measure is calculated as the probability of the preposition given that the verb is present, minus the probability of the preposition given that the verb is absent. Under the predictability hypothesis we would expect to find the shorter expressions *a* and *de* to be more predictable than longer ones *para* and *desde*.

We find that overall, in accordance with the goal-source asymmetry, *a* ‘to’ was the most predictable of the four prepositions, but *desde* ‘from’ was much more predictable than *de* ‘from, of’. Additionally, although *de* is more frequent than *a* (144487746 and 54479710 hits respectively), verb preposition pairs were more frequent with *a* than with *de*. Thus, we find no correlation between length of the expression and predictability of the expression, but we do find a correlation between the raw frequencies of the expressions and their length. An important exception was *salir* which was the only verb that attracted *de* more than *a* ( $\Delta p$  0.17 vs 0.11, respectively). Since *salir* is strongly conceptualized as a source motion verb, this attraction asymmetry speaks for an iconicity effect. If frequency plays a role in marking asymmetries, this is not linked to predictability.

directional		neutral		directional		neutral	
	$\Delta p$		$\Delta p$		mean(freq)		mean(freq)
<i>a</i>	0.2701	<i>a</i>	0.0707	<i>a</i>	656804.143	<i>a</i>	21511.400
<i>de</i>	-0.0009	<i>de</i>	-0.0009	<i>de</i>	85348.857	<i>de</i>	12638.400
<i>desde</i>	0.0034	<i>desde</i>	0.0034	<i>desde</i>	8590.429	<i>desde</i>	1015.000
<i>para</i>	-0.0021	<i>para</i>	-0.0008	<i>para</i>	9992.857	<i>para</i>	1776.400

**References:** • Davies, M. (2015–2016): *Corpus del Español: 2 billion words*. Available online at [www.corpusdelespanol.org](http://www.corpusdelespanol.org). • Davies, M. and M. Ferreira (2015–2016). *Corpus do Português: 1 billion words*. Available online at <http://www.corpusdoportugues.org>. • Gries, S. T. (2013). 50-something years of

work on collocations. In: *International Journal of Corpus Linguistics* 18(1), 137–165. • Haspelmath, M. (2008). Frequency vs. iconicity in explaining grammatical asymmetries. In: *Cognitive Linguistics* 19(1), 1–33. • Lakusta, L. (2005). *Source and goal asymmetries in non-linguistic motion event representations*. PhD thesis. Baltimore, Maryland: Johns Hopkins University. • Lakusta, L. and B. Landau (2005). Starting at the end: the importance of goals in spatial language. In: *Cognition* 96, 1–33. • Markovskaya, E. (2006). Goal-Source Asymmetry and Russian Spatial Prefixes. In: *Tromsø Working Papers in Linguistics*. Ed. by P. Svenonius. 33(2), 200–219. • Nam, S. (2004). Goal and source: Asymmetry in their syntax and semantics. Talk given at: 'Workshop on Event Structures', Leipzig University, Leipzig, March 17–19. • Papafragou, A. (2010). Source-goal asymmetries in motion representation: Implications for language production and comprehension. In: *Cognitive Science*, 34(6), 1064–1092. • Regier, T. and M. Zheng (2007). Attention to Endpoints: A Cross-Linguistic Constraint on Spatial Meaning. In: *Cognitive Science* 31, 705–719. • Stefanowitsch, A. and A. Rohde (2004). The goal bias in the encoding of motion events. In: *Studies in Linguistic Motivation*. Ed. by G. Radden and K.-U. Panther. Berlin: Mouton de Gruyter, 249–268.