



Cognitive Science 46 (2022) e13193

© 2022 The Authors. *Cognitive Science* published by Wiley Periodicals LLC on behalf of Cognitive Science Society (CSS).

ISSN: 1551-6709 online

DOI: 10.1111/cogs.13193

# The In-Out Effect in the Perception and Production of Real Words

Jan A. A. Engelen 

*Department of Communication and Cognition, Tilburg School of Humanities and Digital Sciences, Tilburg University*

Received 13 December 2021; received in revised form 30 June 2022; accepted 6 July 2022

---

## Abstract

The in-out effect refers to the tendency that novel words whose consonants follow an inward-wandering pattern (e.g., P-T-K) are rated more positively than stimuli whose consonants follow an outward-wandering pattern (e.g., K-T-P). While this effect appears to be reliable, it is not yet clear to what extent it generalizes to existing words in a language. In two large-scale studies, we sought to extend the in-out effect from pseudowords to real words and from perception to production. In Study 1, we investigated whether previously collected affective ratings for English and Dutch words were more positive for inward-wandering words and more negative for outward-wandering words. No systematic relationship between wandering direction and affective valence was found. In Study 2, we investigated whether inward-wandering words are more likely to occur in positive online consumer restaurant reviews written in English and Dutch, compared to negative reviews, and whether this association was stronger for food ratings than for decor ratings. Again, no systematic relationship between wandering direction and review rating emerged. We suggest that the affective states triggered by different consonantal wandering directions might be used as a cue for forming judgments in the absence of other information, but that wandering direction is too low in salience to drive the shape of words in the lexicon.

*Keywords:* Language; Affect; Articulation; Iconicity; Word norms; Online reviews

---

---

Correspondence should be sent to Jan Engelen, Department of Communication and Cognition, Warandelaan 2, 5037 AB, Tilburg, The Netherlands. E-mail: J.A.A.Engelen@tilburguniversity.edu

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

## 1. The in-out effect in the perception and production of real words

The in-out effect, first attested by Topolinski, Maschmann, Pecher, and Winkielman (2014), refers to the phenomenon that words whose consonants follow an inward-wandering pattern (e.g., PASOKI, whose consonantal stricture spots are the lips (P), the alveolar ridge (S), and the velum (K), respectively, thus moving from the front of the mouth to the back of the mouth) are preferred over words whose consonants follow an outward-wandering pattern (e.g., KASOPI). It occurs regardless of whether the word is presented as a label for a food item, a dish, a food company, a name for an online marketplace seller, a chat partner, or simply a nonsense word (Silva & Topolinski, 2018; Topolinski & Boecker, 2016a; Topolinski et al., 2014). Furthermore, the effect is robust to variations in whether the word is read silently, read aloud, or played over headphones, whether the participants are native speakers of German, English, French, Portuguese, Turkish, or Ukrainian, and whether the judgment that is made is one of preference, likeability, trustworthiness, or willingness to pay (Garrido, Godinho, & Semin, 2019; Godinho & Garrido, 2016; Godinho, Garrido, & Horchak, 2019; Rossi, Pantoja, Borges, & Werle, 2017; Topolinski, Zürn, & Schneider, 2015, Topolinski & Boecker, 2016b).

Three explanations have been proposed for the in-out effect. First, the articulation of inward and outward words and the ingestion and expectoration of food rely on similar patterns of contraction of oral muscles (Topolinski et al., 2014). Because of this shared functionality, the articulation of inward words may trigger the motivational state of approach that is associated with deglutition; conversely, the articulation of outward words may trigger the motivational state of avoidance associated with expectoration. These motivational states, in turn, influence participants' evaluations of the stimuli they are asked to judge. While there is some supporting evidence (e.g., Rossi et al., 2017), recent findings call this explanation into question. For instance, the in-out effect still occurs when covert articulation is prevented (Lindau & Topolinski, 2018), and neither food deprivation (which should lead to a stronger preference for inward words) nor disgust induction (which should reduce the preference) seems to modulate it (Maschmann, Körner, Boecker, & Topolinski, 2020).

Second, it has been noted that inward words are more frequent in English and German than outward words (Bakhtiari, Körner, & Topolinski, 2016). As a result, adult language users have more experience perceiving and producing inward words. This leads to greater fluency in processing stimuli that follow this pattern, which in turn positively influences language users' evaluations of these stimuli (Reber, Schwarz, & Winkielman, 2004). In line with this, implicitly training participants with massive exposure to (and thus temporarily increasing the fluency of processing of) inward words enhanced the preference for inward words, whereas exposure to outward words reversed the canonical in-out effect and induced a preference for outward words (Körner, Bakhtiari, & Topolinski, 2019). Again, however, empirical support is not unequivocal: the in-out effect does not become stronger in within-subject designs compared to between-subjects designs, as other effects based on fluency tend to do (Godinho & Garrido, 2021), and the training effects mentioned above only seem to extend to words that use the same consonants rather than any inward- or outward-wandering sequence (Ingendahl, Schöne, Wänke, & Vogel, 2021).

Third, Maschmann et al. (2020) argued that the in-out effect may be an epiphenomenon of two more general tendencies: a preference for front consonants over back consonants, and a primacy effect, such that initial consonants play a larger role in the evaluation of a word-like stimulus than later consonants. Research on the affective iconicity of phonemes in the lexicons of various languages seems to partially support this explanation. Adelman, Estes, and Cossu (2018) found that word-initial phonemes, more than subsequent phonemes, predicted the perceived emotional valence of words (e.g., word-initial /p/ was predictive of positive valence, whereas /d/ was predictive of negative valence). However, this relationship was stronger for individual phonemes than for more general phonetic features (e.g., place of articulation). Furthermore, recent findings suggest that front-middle-rear words (e.g., BODIKA) are preferred over front-middle-front words (e.g., BODIBA) (Ingendahl & Vogel, 2022) and that position-weighted consonant preferences apply to consonant sequences presented as fragments, but not to fully spelled-out words (Körner & Rummer, 2022).

In short, the in-out effect has been demonstrated in various task contexts and may rest on one or more well-established cognitive mechanisms, although current evidence is inconclusive about the contribution of each of these mechanisms (see Ingendahl, Vogel, & Topolinski, 2022, for a recent overview). An equally important question, which has received far less attention, is to what extent the in-out effect generalizes to real words. Across all studies on the in-out effect, researchers have relied on pseudowords, such as KENOBA or OGIP, as stimuli. In nearly all cases, moreover, these pseudowords have eschewed consonant clusters, such that subsequent consonants were always separated by a vowel. Thus, it is unclear whether the preference for inward-wandering trajectories is merely a useful insight for marketers who get to coin a name for a new product, or whether it fundamentally affects how speakers of a language feel when they speak, hear, and read words in their language. In this study, we look for the in-out effect “in the wild” in two places: in the lexicons of Dutch and English and in online reviews written in these languages.

### 1.1. *The in-out effect in the lexicon*

We first turn to the possibility of the existence of an in-out effect in the lexicon. While arbitrary form-meaning pairings are a hallmark of natural languages, languages across the world also contain forms which bear a direct resemblance to their meaning (Wichmann, Holman, & Brown, 2010). This phenomenon is known as *iconicity* (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Lupyan & Winter, 2018). An example is onomatopoeia, which refers to lexicalized forms that designate a particular sound, such as *buzz* or *cuckoo*. The in-out effect could be viewed as a particular case of iconicity, such that the ordering of consonantal stricture spots in a word resembles some feature of the concept it denotes. This possibility was already suggested by Topolinski et al. (2014), referring to it as “onomatokinesia” (p. 893). The question of what conceptual features might be involved in that resemblance, however, has not yet been explored in detail.

Here, we make a distinction between semantic and affective meaning. A resemblance at the semantic level would entail the use of inward-wandering consonant sequences in words for ingestion-related concepts (e.g., *drink*) and the use of outward-wandering consonant

sequences in words for expectoration-related concepts (e.g., *cough*). Intriguing as this possibility is, it would only potentially affect a relatively small and circumscribed lexical domain. On the other hand, a resemblance at the affective level would be more indirect, implying an association between inward-wandering consonant sequences and any affectively positive concept, and between outward-wandering consonant sequences and any affectively negative concept. As such, it could be present across the vocabulary of a language and even play a role in shaping words for abstract concepts.

The notion that sublexical components of words may be indicative of their affective meaning is not new. Whissell (1999) found that certain phonemes, such as /l/, across various text genres, occur more often in samples that are rated as pleasant, soft, and passive, whereas others, such as /r/, occur more often in samples that are rated as unpleasant and active. A potential explanation for these patterns is the shared role of muscular effectors for vocalization and emotion. For example, pronouncing /i:/ requires a pattern of muscular contractions similar to smiling, which is associated with happiness; conversely, sounds that are articulated towards the back of the throat, such as /k/ or /g/, require muscular contractions similar to those associated with anger and disgust (Whissell, 1999).

For the German language, Aryani, Conrad, Schmidtke, and Jacobs (2018) calculated a word's "phonological affective potential" (PAP) as a way of quantifying the contribution of a word's sound to a word's affective rating, over and above its semantic content. They then regressed the PAP values on a set of acoustic features (e.g., fundamental frequency, sound intensity). It was found, among other things, that short vowels, voiceless consonants, and hissing sibilants were associated with high arousal and negative valence. This suggests that a word's affective meaning is codetermined by its acoustic-phonetic features. Beyond English and German, Adelman et al. (2018) found that phonological properties significantly predicted valence and arousal ratings for words provided by participants in earlier word norming studies for Spanish, Polish, and Dutch. Specifically, phonemes that can be pronounced rapidly tend to occur at the beginning of negative words, whereas phonemes that take longer to pronounce are found at the beginning of positive words. Presumably, this gives languages an advantage in communicating about dangers and other negative stimuli.

Thus, various articulatory properties of words have been empirically and theoretically linked to their affective meaning. What sets the in-out effect apart from the form-meaning mappings discussed above, is that it is not based on a preference for individual phonemes, but rather arises from the distribution of an articulatory feature over a sequence of phonemes (Ingendahl et al., 2022). In Study 1, we investigate whether a word's wandering direction is related to the affective response it elicits. Conveniently, researchers have already collected affective ratings for large sets of words. For English, Warriner, Kuperman, and Brysbaert (2013) obtained ratings of valence, arousal, and dominance for 13,915 lemmas. In particular, valence ratings were obtained by asking participants to indicate how they felt (from "happy" to "unhappy" on a 9-point scale) while reading each word. For Dutch, Moors et al. (2013) obtained ratings for the same constructs for 4,299 words. Valence ratings were obtained by asking participants to judge whether the word referred to something "positive/pleasant" or "negative/unpleasant" on a 7-point scale. These valence ratings are similar, albeit not identical, to the type of judgments (which, as mentioned, cover a broad

spectrum) that are susceptible to the experimental manipulation of consonantal wandering direction. If variance in these valence ratings is explained by the words' consonantal wandering directions, that would be compelling evidence for the pervasiveness of the in-out effect.

### 1.2. *The in-out effect in online reviews*

Another area where the in-out effect might manifest itself is discourse. Even if there would be no meaningful in-out effect in the lexicon, when people produce spoken or written discourse, they might still choose “from among those words available, the words best matching an ongoing mood” (Whissell, 1999, p. 45). The link between affect and language production is fairly well documented. For example, after watching a video clip, people in a positive mood give more abstract-level descriptions of the events it contained, whereas people in a negative mood use more concrete descriptions (Beukeboom & Semin, 2006). In Study 2, we ask whether a language user's affective state is associated with the frequency of use of inward versus outward words. We do so by analyzing online restaurant reviews. While this is a somewhat specialized genre that should not be taken to be representative of discourse in general, it offers two key advantages.

First, most online review systems combine text with a numerical rating. The valence of the review is directly accessible in the form of this rating. This rating is provided by the reviewer themselves, rather than by an external judge. Moreover, the process of writing a restaurant review shares characteristics with the mood induction procedure of *imagination*. In an imagination mood induction procedure, participants are asked to write down an experience from their life that elicited a particular mood, along with the perceptions, sensations, and affective reactions that were part of that experience (Westermann, Spies, Stahl, & Hesse, 1996). We contend that when a person writes a restaurant review, they do something similar. They recall and selectively write down the details of a particular episode (e.g., when the visit took place, what food and drinks were consumed) as well as the perceptions and sensations (e.g., how the food looked, smelled, and tasted) and affective reactions they experienced during this episode (e.g., whether the food was delightful or disappointing, whether waiting staff was amicable or hostile). Thus, there is good reason to assume that the review rating serves as a proxy for the writer's affective state.

A second advantage is that restaurant reviews, in contrast to reviews of other types of experience goods, are explicitly concerned with food consumption. Some review websites even give users the possibility to rate the food separately from other aspects of the restaurant. We capitalize on this feature to investigate the much-researched, but a contentious deglutition-expectoration account of the in-out effect (Ingendahl & Vogel, 2022; Maschmann et al., 2020) in a naturalistic context. If the in-out effect is grounded in the biomechanics of swallowing and spitting out, we predict the ratio of inward versus outward words to correlate more strongly with the rating of the food, which bears directly on the degustational pleasure or disgust that was experienced, than with the rating of a more remote component, such as the restaurant's decor.

### 1.3. Overview of studies

We investigated the relationship between consonantal wandering direction and affective valence in the lexicon and in discourse in two sets of studies. An overview of the data included in each study is provided in Table 1. In Study 1, we tested whether previously collected valence ratings for Dutch and English words can be explained by the wandering direction of these words. For English, we looked at two sets of words: words that fit the broad *theoretical* constraints that have been defined for the in-out effect and those that fit the narrower constraints that have been *empirically* demonstrated to produce the in-out effect. In Study 2a, we tested whether the prevalence of inward versus outward words in Dutch online restaurant reviews is correlated with the numerical ratings for food and decor included in these reviews. In Study 2b, we extend this analysis to a larger set of reviews in Dutch and English, using only the overall review rating. In Study 2c, finally, we tested whether words that are presented as stand-alone sentences in reviews are more likely to be inward or outward, depending on the review rating.

## 2. Study 1: The in-out effect in the Dutch and English lexicons

For existing Dutch and English affective word ratings, we tested to what extent variance in valence ratings was explained by a word's consonantal wandering direction. We used phonological transcriptions of words, rather than their spelling, as a basis for determining their consonantal wandering patterns. An advantage of this is that we could include words containing orthographic symbols that could have different pronunciations in different contexts (such as *g* in *goal* vs. *giant*). For English, we conducted our analyses on a subset of words that fit the constraints of stimulus pool D from Topolinski et al. (2014), which was used in the first demonstration of the in-out effect among English speakers, as well as a set of words including a more diverse set of consonants.

### 2.1. Method

#### 2.1.1. Sample

*Dutch words* Affective ratings for 4,299 Dutch words were extracted from Moors et al. (2013). For each word, the orthographic word form was looked up in a list of 300,445 Dutch word forms with matching phonological transcriptions in SAMPA notation from the CELEX database (Baayen, Piepenbrock, & Van Rijn, 1993). The SAMPA symbols *p*, *b*, *m*, *f*, and *v* (representing bilabial and labiodental phonemes) were considered front consonants, *t*, *d*, *n*, *s*, *z*, *ʃ*, *ʒ*, and *l* (representing alveolar and post-alveolar) middle consonants, and *k*, *g*, *ŋ*, *x*, *ɣ*, and *h* (representing velar, uvular, and glottal phonemes) back consonants. Because the in-out effect has been explicitly defined for consonants and not vowels (e.g., Topolinski & Boecker, 2016b), words starting or ending with the semivowels *j* or *w* were excluded.

Phonetic transcriptions corresponded to Standard Dutch. This means that word-final *n*, when preceded by schwa, was silent, and thus not part of the consonants that were analyzed. This also means that not every individual might pronounce the words exactly as transcribed.

Table 1  
Overview of data included in each study

Study	Subset	Units of Analysis	Number of Units	In-Out Scale	Valence Scale	Data Source
1	Dutch	Individual words	3,232	-1 (outward), 0 (neutral), 1 (inward) <sup>a</sup>	Valence rating (1-5)	Moors et al. (2013)
	English 1	Individual words	10,408	-1 (outward), 0 (neutral), 1 (inward)	Valence rating (1-7)	Warriner et al. (2013)
	English 2	Individual words	3,420	-1 (outward), 0 (neutral), 1 (inward)	Valence rating (1-7)	Warriner et al. (2013)
2a	Dutch	Aggregates of five words	5,000	-1 (all outward) to 1 (all inward)	Food rating (1-10), Decor rating (1-10)	Iens corpus
2b	Dutch	Aggregates of 20 words	10,000	-1 (all outward) to 1 (all inward)	Review rating (1-10)	Iens corpus
	English	Aggregates of 20 words	10,000	-1 (all outward) to 1 (all inward)	Review rating (1-5)	Yelp dataset
2c	Dutch	One-word sentences	1,000	-1 (outward), 0 (neutral), 1 (inward)	review rating (1-10)	Iens corpus
	English	One-word sentences	1,000	-1 (outward), 0 (neutral), 1 (inward)	review rating (1-5)	Yelp dataset

<sup>a</sup>For Study 1, several alternative operationalizations of inward and outward wandering trajectories were investigated besides the canonical ones reported in the paper. These are available on OSF (<https://osf.io/6wvyc/>).

With respect to consonants, there are considerable regional variations in voicing of fricatives (e.g., *s* vs. *z*) and palatalization of alveolar fricatives (e.g., *ʃ* vs. *s*), but none of these affect the general place of articulation that is of relevance for the current study. The only exception is *r*, which has allophones that do not allow it to be unambiguously classified as middle or back. Therefore, words with *r* as the first or last consonant were excluded from analysis.

In total, 4,263 (99.2%) words could be automatically matched with a phonological transcription. Of these, 67 words contained fewer than two consonants, and 964 words contained *r*, *j*, or *w* as the first or last consonant, meaning that 3,232 words remained for further analysis.

*English words* Affective ratings for 13,915 English words were extracted from Warriner et al. (2013). We first removed 131 items that included a space (e.g., compound nouns such as *action figure*, adjective–noun combinations such as *green bean*, and proper nouns such as *El Niño*) or for which no frequency information could be retrieved from the Subtlex-US corpus (Brysbaert & New, 2009). Then, for each word, the orthographic form was looked up in the Carnegie Mellon Pronouncing Dictionary (version 7b), which contains phonetic transcriptions in ARPAbet for over 1.3 million word forms (Carnegie Mellon University Speech Group, 2014). For words that were not in this list, a rule-based pronunciation was generated with the Logos Lexicon Tool (Carnegie Mellon University Speech Group, 2008). The phonological transcriptions match North American pronunciations. Given that the participants in Warriner et al. (2013) that rated the words were U.S. citizens, this is appropriate. Still, it is worth pointing out that not every individual might pronounce a given word the same way. For instance, the suffix *-ing* was transcribed as *IHNG*, thus ending with a velar nasal, which was classified as a back consonant.

*Word set 1* The ARPAbet symbols *P*, *B*, *M*, *F*, *V*, *TH*, and *DH* (representing bilabial, labiodental, and dental phonemes; thus any consonants involving the lips and teeth) were considered front consonants, *T*, *D*, *S*, *Z*, *N*, *CH*, *SH*, *ZH*, *L*, and *JH* (representing alveolar, post-alveolar, and palatal phonemes) middle consonants, and *K*, *G*, *NG*, and *HH* (representing velar and glottal phonemes) back consonants. As in the Dutch dataset, words that included *R* or one of the semivowels *W* and *Y* as its first or last consonant were excluded. (Note that a word's first consonant need not be word- or syllable-initial, as in *army*.) In total, 177 words contained fewer than two consonants, and 3,199 words contained the consonant *R*, *W*, or *Y* as its first or last consonant, meaning that 10,408 words remained for further analysis.

*Word set 2* To stay close to the original demonstration of the in-out effect with speakers of English, we also analyzed the English data using only the consonants that were used in words for English-speaking participants in Pool D from Topolinski et al. (2014). These were ARPAbet symbols *B*, *F*, *M*, and *P* as front consonants, *D*, *L*, *N*, *S*, and *T* as middle consonants, and *K* as the only back consonant. Words that had any other consonant in them, regardless of where in the word these occurred, were discarded from analysis. We did, however, not require subsequent consonants to be separated by a vowel, and we did not require words to contain exactly three consonants, as there were only 55 words in the dataset that fit both these criteria



(e.g., *family*, *banana*, *casino*). By relaxing these two criteria, there were 3,420 English words that could be analyzed as part of Word set 2.

### 2.1.2. Data analysis

For both Dutch and English words, the phonetic transcriptions were used to create arrays in which each consonant was represented with an integer: 1 for front consonants, 2 for middle consonants, and 3 for rear consonants. The overall direction of a word was determined by subtracting the last integer from the first integer and then converting any positive value to 1 and any negative value to  $-1$ . This resulted in values of 1 for inward words,  $-1$  for outward words, and 0 for words whose first and last consonant shared the same general place of articulation.<sup>1</sup>

For each set of words, Pearson correlations between valence ratings and consonantal wandering directions were computed. To maximize comparability with previous studies on the in-out effect, in which the stimuli to be judged were typically presented as objects, rather than actions, events, or attributes, and could hence be seen as nouns, we conducted these correlations separately for nouns and for all words together. In addition, to investigate how much unique variance is explained by word direction after controlling for several other lexico-semantic predictors that are known to predict valence ratings (Citron, Weekes, & Ferstl, 2014), a forward stepwise linear regression analysis was conducted on the valence ratings for all words, using orthographic word length, number of phonemes, log-10 word frequency (for Dutch words, the frequency was based on the Subtlex-NL corpus (Keuleers, Brysbaert, & New, 2010) and already included in the dataset; for the English words, these were based on the Subtlex-US corpus (Brysbaert & New, 2009) and retrieved separately), arousal rating, dominance rating, and consonantal wandering direction as predictors.<sup>2</sup> At each step, predictors were selected based on  $p$ -values, such that a predictor was only included in the model if it was significant at  $\alpha < .05$ .

## 2.2. Results

Table 2 shows the frequency and mean valence rating of each articulation pattern in the dataset of Dutch and English word ratings. In each word set, inward words were more common than outward words. Furthermore, in each word set, the mean valence rating was slightly higher than the midpoint of the scale, indicating that the words were on average rated as neutral to mildly pleasant. However, differences between articulation patterns were numerically very small.

For Dutch, there was no significant zero-order correlation between consonantal wandering direction and valence rating for nouns ( $N = 1928$ ),  $r = -.014$ ,  $p = .531$ , nor for all words ( $N = 3232$ ),  $r = -.012$ ,  $p = .502$ . The regression model that accounted for most of the variance (13.9%) in valence ratings for all words included the predictors dominance rating, arousal rating, log-10 frequency, and number of phonemes. Adding consonantal wandering direction did not significantly improve the fit of the model. Table 3 shows the regression model for the valence ratings of Dutch words.

Table 2

Frequencies and valence ratings for inward, neutral, and outward words in Dutch and English

Subset	Direction	Dutch			English 1			English 2		
		<i>n</i>	Valence		<i>n</i>	Valence		<i>n</i>	Valence	
			<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>
Nouns	Inward (1)	715	3.96	0.98	2,248	5.17	1.22	750	5.15	1.21
	Neutral (0)	645	3.97	1.02	2,473	5.06	1.25	889	5.1	1.24
	Outward (-1)	568	3.99	0.91	1,697	5.11	1.27	481	5.19	1.23
All words	Inward (1)	1,270	3.91	1.10	3,695	5.08	1.29	1,217	5.06	1.26
	Neutral (0)	1,080	3.93	1.08	4,159	4.99	1.30	1,499	5.02	1.26
	Outward (-1)	882	3.94	1.00	2,554	5.08	1.31	704	5.15	1.24

*Note.* Valence ratings for English were provided on a 1–9 scale; valence ratings for Dutch were provided on a 1–7 scale.

Table 3

Regression model for Dutch word valence ratings

Predictor	<i>b</i>	<i>SE</i>	$\beta$	<i>t</i>	<i>p</i>
(Intercept)	2.544	0.114		22.378	<.001
Dominance	0.668	0.032	0.434	20.992	<.001
Arousal	-0.318	0.026	-0.252	-12.175	<.001
Log-10 frequency	0.099	0.021	0.081	4.625	<.001
Phonemes	-0.027	0.009	-0.051	-2.914	.004

For English word set 1, there was no significant zero-order correlation between consonantal wandering direction and valence rating, for nouns ( $N = 6,418$ ),  $r = .021$ ,  $p = .086$ , nor for all words ( $N = 10,408$ ),  $r = .004$ ,  $p = .702$ . The regression model that accounted for most of the variance (53.7%) in valence ratings for all words included the predictors dominance rating, word frequency, arousal rating, and orthographic length. Consonantal wandering direction did not significantly improve the fit of the model. Table 4 shows the regression model for English word set 1.

Table 4

Regression model for English word valence ratings (Set 1)

Predictor	<i>b</i>	<i>SE</i>	$\beta$	<i>t</i>	<i>p</i>
(Intercept)	-0.025	0.077		-0.330	.742
Dominance	0.960	0.009	0.706	102.612	<.001
Log-10 frequency	0.158	0.014	0.081	11.183	<.001
Arousal	-0.083	0.010	-0.057	-8.325	<.001
Orthographic length	0.017	0.004	0.032	4.438	<.001

Table 5  
Regression model for English word valence ratings (Set 2)

Predictor	<i>b</i>	<i>SE</i>	$\beta$	<i>t</i>	<i>p</i>
(Intercept)	0.040	0.135		0.293	.770
Dominance	0.935	0.017	0.687	55.781	<.001
Log-10 frequency	0.150	0.023	0.086	6.521	<.001
Arousal	−0.075	0.017	−0.053	−4.310	<.001
Orthographic length	0.025	0.008	0.043	3.250	.001

For English word set 2, there was no significant zero-order correlation between consonantal wandering direction and valence rating for nouns ( $N = 2,120$ ),  $r = -.010$ ,  $p = .652$ , nor for all words ( $N = 3,420$ ),  $r = -.022$ ,  $p = .199$ . The regression model that accounted for most of the variance (50.8%) in valence ratings for all words included the predictors dominance rating, log-10 frequency, arousal rating, and orthographic length. Consonantal wandering direction did not significantly improve the fit of the model. Table 5 shows the regression model for English word set 2.

### 3. Study 2a: The in-out effect in Dutch online reviews

To analyze the relationship between affective valence and the prevalence of inward and outward words at the discourse level, we used the Iens corpus (Engelen & Kraemer, 2021), a collection of over 684,000 Dutch consumer reviews of restaurants. A distinguishing feature of this corpus is that reviews contain separate ratings for food, service, and decor. The frequency of inward and outward words was estimated by drawing samples of words from a large number of reviews, obtaining phonological transcriptions of these words, and calculating the average wandering direction for each sample. As in Study 1, to keep our study close to earlier work on the in-out effect, in which the stimuli to be judged were typically presented as referring to objects, we focused on nouns.

It should be noted that this method represents a rather simplified approach to discourse analysis. By ignoring negations (e.g., “The vegetables were *not* well seasoned”) and quantifiers (e.g., “The home-made mayonnaise was *extremely* tasty”), but also the syntactic roles of words and the order in which they occur, this approach implicitly assumes that the valence of a text is a linear combination of the valence of its component words. For sentences, however, affective meaning is better predicted by an approach that takes supralexical features into account (Lüdtke & Jacobs, 2015). Nevertheless, while not performing as well, even simple dictionary methods that consider only the valence of random samples of individual words (e.g., Bestgen, 1994) account for considerable variance in valence ratings for sentences and texts. In line with this, the number of negative emotion words in restaurant reviews significantly predicts review ratings (Jurafsky, Chahuneau, Routledge, & Smith, 2014). Therefore, while we do not expect a very strong correlation, the ratings of reviews should, in principle, still be predictable from properties of individual words that are sampled from them.

### 3.1. Method

#### 3.1.1. Sample

We randomly sampled 5,000 reviews from the Iens corpus, such that for each overall rating (derived by rounding the arithmetic mean of the three partial grades to the nearest integer) there were 500 reviews. Further constraints were that no more than one review was selected for a particular reviewer and that each review contained at least five nouns whose phonological transcription contained at least two consonants (excluding those with *r*, *j*, or *w* as the first or last consonant). The reviews were tokenized and Part-of-Speech (PoS) tagged with the spaCy library (version 2.3) (Honnibal & Montani, 2017) for Python, using the small-sized language model for Dutch, which has a PoS tagging accuracy of about 97%. Nouns comprised 16% of all word tokens in the corpus.

#### 3.1.2. Data analysis

The same classification for front, middle, and back consonants was used as in Study 1. Inward, outward, and neutral words were scored as 1,  $-1$ , and 0, respectively. Wandering directions for the five nouns were averaged for each review, yielding a score between  $-1$  and 1, with 0 indicating a balance between inward and outward words. We computed Pearson correlations between averaged wandering directions and ratings for food and decor (both on a 1–10 scale, with increments of 1), respectively.

### 3.2. Results

Of the 25,000 nouns sampled from the reviews, 11,736 (46.9%) were inward and 6,526 (26.1%) were outward. Wandering direction was not significantly correlated with food rating,  $r = -.018$ ,  $p = .216$  nor with decor rating,  $r = -.001$ ,  $p = .921$ . Thus, the data do not provide evidence for an in-out effect in Dutch online reviews.

## 4. Study 2b: The in-out effect in Dutch and English online reviews

By focusing solely on nouns, our results may not have been representative of the distribution of *all* inward and outward wandering words in the corpus. With respect to the perceived iconicity of various parts of speech, nouns are somewhere in the middle: less iconic than adjectives, verbs, and interjections, but more iconic than names and function words (Winter, Perlman, Perry, & Lupyan, 2017). By looking at all parts of speech together, then, there is a possibility that an in-out effect would be “drowned out” by a large proportion of highly frequent function words (pronouns, demonstratives, prepositions, conjunctions, etc.), which, following from a Zipfian distribution, would inevitably dominate the sample. On the other hand, even usage patterns of function words have been shown to correlate with a writer’s mood (Rude, Gortner, & Pennebaker, 2004; Weintraub, 1989). Being in a more positive or negative affective state, therefore, could potentially also influence the distribution of inward versus outward words among parts of speech that are generally perceived as less iconic. In Study 2b, therefore, we took a larger and less restrictive sample of reviews and words within a

review, including all parts of speech. Furthermore, to decrease the likelihood that the findings were particular to the Dutch language or culture, we performed similar analyses on English restaurant reviews from the Yelp dataset (Yelp Inc., 2019), which contains reviews from several North American metropolitan areas.

#### 4.1. Method

##### 4.1.1. Sample

For the Dutch Iens corpus, the analysis procedure was identical to Study 1a, except that we sampled 20 words each from 10,000 reviews without any constraints on the parts of speech and without setting a quota for rating categories. The reason we required each review to contribute a fixed number of words was to avoid heteroskedasticity, as short reviews (e.g., 20 words) can be expected to show more extreme scores, and therefore larger error variances, than longer reviews (e.g., 200 words).

For the Yelp dataset, we extracted English restaurant reviews from the *Review* part for users who were also in the *User* part. This resulted in a set of 192,025 reviews, which each contain a general rating between 1 and 5 stars (with increments of 0.5 stars, making this essentially a 9-point scale). The reviews were tokenized with spaCy (version 3.0), using the medium-sized language model for English. Reviews and words within reviews were sampled using the same procedure as for the Dutch corpus. Because the Yelp dataset did not contain separate food ratings, we used the overall review ratings in the analysis of both datasets.

##### 4.1.2. Data analysis

For Dutch, we used the same classification for inward, outward, and neutral words as in Study 1 and Study 2a; for English, we used the classification for word set 1 in Study 1. Wandering directions for the 20 words were averaged for each review, yielding a score between  $-1$  and  $1$ . Pearson correlations between averaged wandering direction and overall review rating (ranging from 1 to 10 with increments of 0.1 for Dutch and from 1 to 5 with increments of 0.5 for English) were computed for both languages.

#### 4.2. Results

Of the 200,000 words sampled from the Dutch Iens corpus, 77,303 (38.7%) were inward and 53,735 (26.9%) were outward. There was no correlation between review rating and wandering direction,  $r = -.019$ ,  $p = .063$ . Of the 200,000 words sampled from the Yelp corpus, 76,004 (38.0%) were inward and 45,743 (22.7%) were outward.<sup>3</sup> There was no correlation between review rating and wandering direction ( $r = -.004$ ,  $p = .701$ ). Fig. 1 (top row) shows the distribution of review wandering directions by rating for both datasets.

### 5. Study 2c: The in-out effect in one-word sentences

A possible explanation for the lack of evidence for an in-out effect in language production lies in the fact that in perception studies, stimuli are presented as isolated words, making

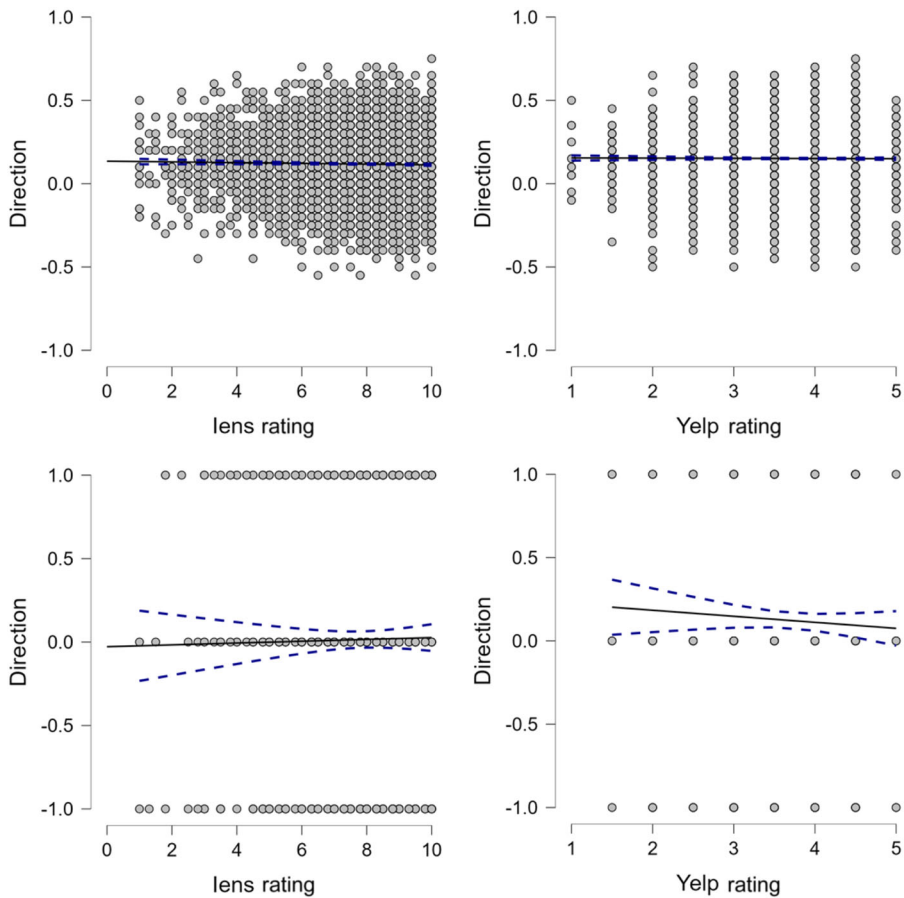


Fig. 1. Wandering direction by review rating.

Note. Top row: random samples of 20 words per review for Dutch (left panel) and English (right panel) from Study 2b. Bottom row: one-word utterances from reviews in Dutch (left panel) and English (right panel) from Study 2c. The superimposed lines represent the regression lines with 95% confidence intervals.

word boundaries more salient than they are in longer utterances. In connected speech, words typically bleed into one another. For instance, coarticulation straddles word boundaries (e.g., *lean bacon* is pronounced as “leam bacon”) (Gaskell & Marslen-Wilson, 1996), and there is evidence that the phonological encoding stage of speech planning involves so-called phonological words, which are representations that may span multiple lexical words (such as “de-man-dit” in *police demand it*) (Levelt & Wheeldon, 1994). While these findings are specific for production of spoken language, similar effects may occur when language is written or typed.

To investigate the role of salience of word boundaries, we conducted an analysis on one-word sentences embedded in Dutch and English reviews (e.g., “Delicious!” or “Gross.”). The

rationale is that these words would typically be preceded and followed by a distinct pause if the review were read aloud, thus making their boundaries more salient. Although one-word sentences are not that frequent in coherent written discourse and only a limited set of words can be used in such utterances, the size of the corpora allowed us to retrieve large enough samples. We predicted one-word sentences embedded in positive reviews to show a stronger tendency to be inward-wandering than those in negative reviews.

### 5.1. Method

For English, one-word sentences were extracted by splitting the reviews into sentences using the sentencizer module in spaCy (version 3.0) and then extracting sentences consisting of a single word token. For Dutch, results from the sentencizer module were not satisfactory, so one-word utterances were extracted with a regular expression pattern that matched any sequence of a sentence-ending punctuation characters (“.” “!” and “?”), followed by one or more whitespace characters, followed by one or more alphanumeric characters, followed by a sentence-ending punctuation character. From both the Dutch and English datasets, we sampled 1,000 reviews that had at least one one-word sentence. From reviews that had multiple such utterances, we randomly sampled one. Further constraints were that no more than one review per user was sampled.

### 5.2. Results

Of the 1,000 one-word sentences sampled from the Dutch lens corpus, 318 were inward and 303 were outward. There was no correlation between review rating and wandering direction,  $r = .011$ ,  $p = .718$ . Of the 1,000 one-word sentences sampled from the English Yelp corpus, 342 were inward and 220 were outward. There was no correlation between review rating and wandering direction,  $r = -.031$ ,  $p = .323$ . Fig. 1 (bottom row) shows the distribution of wandering directions by rating for both datasets.

## 6. General discussion

We looked for the in-out effect outside of the lab and beyond highly constrained sets of pseudowords by investigating the association between consonantal wandering direction at the word level on the one hand and valence at the word (Study 1) and discourse level (Study 2) on the other. Across two languages, Dutch and English, using very large samples and using both liberal and conservative operationalizations of inward and outward words, we did not find any of the hypothesized correlations between articulation dynamics and affective valence.

We tried to keep the definitions of inward and outward words as close to earlier work as possible, but some adjustments were inevitable. Even the set in Study 1 that was made up of a very restrictive set of consonants was larger than the set of words in stimulus pool D of Topolinski et al. (2014). Would we have applied those exact criteria, such that only words whose consonants would follow the pattern front-middle-back or back-middle-front, each consonant being separated by a vowel, were included, then the Warriner et al. (2013) dataset

of 13,915 words would have yielded only 11 words. (The exhaustive list is: *academy*, *basic*, *boutique*, *canopy*, *column*, *cutoff*, *economy*, *manic*, *medic*, *panic*, and *panicky*.) This speaks to the limited representativeness of the stimuli in earlier studies for the English language.

An important assumption in research on the in-out effect is that it arises from a word's phonological, rather than orthographical properties (Topolinski & Boecker, 2016b; Topolinski et al., 2014). We, therefore, chose to identify a word's articulation dynamics based on its phonetic transcription, rather than its orthographic form. This way, we could include a large number of words in our analyses which included one or more orthographic symbols that, by themselves, may correspond to more than one sound. For instance, the symbol *g* in written English can be a voiced post-alveolar fricative (as in *beige*) or affricate (as in *giant*), or a voiced velar plosive (as in *goal*). There is a possibility, however, that certain orthographic forms are dispreferred, regardless of their pronunciation. Irregular words (such as *through* or *pint*) or words containing irregular grapheme-phoneme mappings, are read less fluently than regular forms (Seidenberg, Waters, Barnes, & Tanenhaus, 1984) and should therefore, under the fluency account, be rated as less pleasant. Still, even when we analyzed words that used only consonants with highly regular pronunciations (English set 2 in Study 1), no in-out effect emerged.

The exact variable whose relation to consonantal wandering direction we investigated was different across each of our studies. In Study 1, the dependent variables were ratings on a scale from "happy" to "unhappy" for English and from "pleasant" to "unpleasant" for Dutch. In Study 2a, we used review ratings for food and decor, and in Studies 2b and 2c we used overall review ratings. This multiplicity is mirrored in the literature: consonantal wandering direction of a word influences various types of judgments, such as likeability (Topolinski et al., 2014), trustworthiness (Silva & Topolinski, 2018), and perceived palatability (Rossi et al., 2017). While the ratings used in the present study arguably fall under the same conceptual umbrella, it is possible that none of them would be sensitive to wandering direction in more controlled studies. More generally, the variety of judgments for which an in-out effect has been shown is sometimes seen as an indicator of its robustness, but it also highlights the current theoretical underspecification of the underlying psychological mechanism. Future laboratory studies could investigate the boundaries of the types of judgment that are sensitive to a word's wandering direction to better understand exactly what psychological states inward and outward words bring about.

A finding from previous research that we did replicate was the prevalence of inward words in general. At both the type (Study 1) and the token level (Study 2) inward words were about 1.5 times more frequent than outward words, although in English words whose first and last consonant shared the same general place of articulation were even more frequent. These findings go beyond those presented by Bakhtiari et al. (2016) in two ways. First, their estimates did not ostensibly account for differences in how often language users would actually encounter specific words, which we did by looking at token frequencies. Second, their sample of words was restricted by the need for orthographic symbols to have consistent pronunciations, which we circumvented by using phonological transcriptions. Arguably, the prevalence of inward words could make production and perception of these words more fluent, but regardless of their fluency they were not associated with more positive valence ratings.



We suggested that the form-to-meaning mapping underlying the in-out effect might be viewed as a particular case of affective iconicity. Iconicity, in general, has advantages for language learning and communication. For example, iconic signs are recognized more easily and reproduced more accurately than more arbitrary signs (Thompson, Vinson, Woll, & Vigliocco, 2012). However, for iconicity to be useful in communication, we contend that the resemblance between form and meaning, whether semantic or affective, needs to be at the very least *noticeable*. Unlike the sounds that make up a particular word, which are salient to both the speaker and the listener, the articulatory kinematics are experienced primarily by the speaker—although there is a long tradition of research supporting the motor theory of speech perception (see Galantucci, Fowler, & Turvey, 2006, for a review), which postulates that recognizing speech sounds requires covert imitation of speech gestures. Some of the studies on the in-out effect used suspicion checks and post-experimental interviews (e.g., Topolinski et al., 2014), none of which yielded evidence that participants had detected patterns in the words. With consonantal wandering directions going largely unnoticed, there is little reason to assume these give a communicative advantage substantial enough to be used as iconic expressions for affectively positive and negative concepts.

It should be noted that the majority of research on the in-out effect has been conducted with speakers of German, a language that was not considered in the present analyses. There are only a few published studies that demonstrated the in-out effect with English speakers (e.g., Rossi et al., 2017; Topolinski et al., 2014), and, to date, none with Dutch speakers. It can therefore not be ruled out that the Dutch language, even in more controlled studies, simply resists the in-out effect. Given its close typological proximity to both English and German, however, we deem such an explanation unlikely. These languages share a long common history, which shows in similar lexical forms and a large number of cognates. Even historical developments in the phonology of these languages, such as the second Germanic consonant shift (around 700 C.E., Iverson & Salmons, 2006), which sets German apart from English and Dutch, do not seem to impact the wandering direction of words. For example, in certain contexts plosives became fricatives or affricates in German (such as *Apfel*, compared to English *apple* and Dutch *appel*), but their general place of articulation has stayed the same. Thus, if there is no in-out effect in two of these languages, it is unlikely to show up in a third.

By that same token, one might argue Dutch and English do not represent an entirely independent sample of languages. Moreover, as Adelman et al. (2018) point out, sound symbolism is generally weaker in Indo-European languages than in other language families (see also Perniss, Thompson, & Vigliocco, 2010). It might thus be worthwhile to investigate the in-out effect in real words in other languages. These could also include languages with more transparent orthographies to further probe potential orthographical explanations for the present findings. With the proliferation of online reviews for almost any kind of goods, explorations along the lines of Study 2 are possible for other languages. Moreover, associations between valence and articulatory kinematics in the lexicons of other languages could be studied with the help of the National Research Council Canada Valence, Arousal and Dominance (NRC VAD) Lexicon (Mohammad, 2018), which contains ratings for 20,000 words and has been translated from English to over 100 languages.

To sum up, two large-scale studies demonstrate that a reliable, seemingly versatile effect does not readily generalize from highly controlled pseudoword stimuli in the lab to the more messy environments of actual language use. Thus, while the distribution of front, middle, and back consonants over serial positions may serve as a cue for forming affective judgments in tasks where stimuli convey very little other information, it might not be sufficiently salient to tilt affective ratings for words that are already imbued with lexical meanings or to influence the forms that languages use to express these meanings.


## Acknowledgments

I thank Emiel Krahmer and three anonymous reviewers for their insightful comments on this manuscript and Connie de Vos for discussion. Data and analyses presented in this manuscript are available on Open Science Framework (<https://osf.io/6wyce/>).

## Notes

- 1 In exploratory analyses, we also (1) distinguished between “full” (i.e., covering the whole trajectory of front to back or vice versa) and “partial” (i.e., covering only a part of that trajectory) directional patterns, (2) made more fine-grained distinctions along the sagittal axis (cf. Topolinski & Boecker, 2016b) with up to nine articulation spots, and (3) distinguished several other subpatterns, such as whether a word contained any directional reversals (e.g., front-back-middle). These analyses can be found on OSF (<https://osf.io/6wyce/>).
- 2 An assumption underlying both the correlation and regression analyses is that wandering direction is an interval variable. One could debate whether this is the case. We therefore also ran analyses that treated wandering direction as a three-level factor. The parameter estimates of these analyses are very close to the ones reported here and, crucially, word direction was not statistically significant. These analyses are available on OSF (<https://osf.io/6wyce/>).
- 3 It can be empirically shown that it is *words*, rather than arbitrary phoneme sequences in the review texts, that are more likely to be inward- than outward-wandering. These analyses are provided on OSF (<https://osf.io/6wyce/>).

## Open Research Badges

 This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at <https://osf.io/6wyce/>.

## Conflict of interest

I have no competing interests to disclose.

## References

- Adelman, J. S., Estes, Z., & Cossu, M. (2018). Emotional sound symbolism: Languages rapidly signal valence via phonemes. *Cognition*, *175*, 122–130.
- Aryani, A., Conrad, M., Schmidtke, D., & Jacobs, A. (2018). Why 'piss' is ruder than 'pee'? The role of sound in affective meaning making. *PLoS One*, *13*(6), e0198430.
- Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX Lexical Database [CD-ROM]*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.
- Bakhtiari, G., Körner, A., & Topolinski, S. (2016). The role of fluency in preferences for inward over outward words. *Acta Psychologica*, *171*, 110–117.
- Bestgen, Y. (1994). Can emotional valence in stories be determined from words? *Cognition and Emotion*, *8*, 21–36.
- Beukeboom, C. J., & Semin, G. R. (2006). How mood turns on language. *Journal of Experimental Social Psychology*, *42*, 553–566.
- Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*, 977–990.
- Carnegie Mellon University Speech Group. (2008). LOGIOS Lexicon Tool. Retrieved from <http://www.speech.cs.cmu.edu/tools/lextool.html>
- Carnegie Mellon University Speech Group (2014). CMUdict: The Carnegie Mellon Pronouncing Dictionary. Retrieved from <http://www.speech.cs.cmu.edu/cgi-bin/cmudict>
- Citron, F. M., Weekes, B. S., & Ferstl, E. C. (2014). How are affective word ratings related to lexicosemantic properties? Evidence from the Sussex Affective Word List. *Applied Psycholinguistics*, *35*, 313–331.
- Dingemanse, M., Blasi, D. E., Luyyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, *19*, 603–615.
- Engelen, J. A. A., & Krahmer, E. J. (2021, July 9). *Hungry for language data? Introducing a large Dutch corpus of restaurant reviews* [Poster presentation]. Computational Linguistics in The Netherlands 31, Ghent, Belgium.
- Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, *13*, 361–377.
- Garrido, M. V., Godinho, S., & Semin, G. R. (2019). The “ins” and “outs” of person perception: The influence of consonant wanderings in judgments of warmth and competence. *Journal of Experimental Social Psychology*, *82*, 1–5.
- Gaskell, M. G., & Marslen-Wilson, W. D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 144–115.
- Godinho, S., & Garrido, M. V. (2016). Oral approach-avoidance: A replication and extension for European–Portuguese phonation. *European Journal of Social Psychology*, *46*, 260–264.
- Godinho, S., & Garrido, M. V. (2021). The in–out effect: Examining the role of perceptual fluency in the preference for words with inward-wandering consonantal articulation. *Psychological Research*, *85*, 112–120.
- Godinho, S., Garrido, M. V., & Horchak, O. V. (2019). Oral approach avoidance: A replication and extension for Slavic and Turkic phonations. *Experimental Psychology*, *66*, 355–360.
- Honnibal, M., & Montani, I. (2017). *spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing*. Retrieved from <https://spacy.io/>
- Ingendahl, M., Schöne, T., Wänke, M., & Vogel, T. (2021). Fluency in the in-out effect: The role of structural mere exposure effects. *Journal of Experimental Social Psychology*, *92*, 104079.
- Ingendahl, M., & Vogel, T. (2022). The articulatory in-out effect: Driven by consonant preferences? *Journal of Personality and Social Psychology*, *122*(2), e1–e10.
- Ingendahl, M., Vogel, T., & Topolinski, S. (2022). The articulatory in-out effect: Replicable, but inexplicable. *Trends in Cognitive Sciences*, *26*, 8–10.
- Iverson, G., & Salmons, J. (2006). Fundamental Regularities in the Second Consonant Shift. *Journal of Germanic Linguistics*, *18*, 45–70.

- Jurafsky, D., Chahuneau, V., Routledge, B. R., & Smith, N. A. (2014). Narrative framing of consumer sentiment in online restaurant reviews. *First Monday*, 19 <https://doi.org/10.5210/fm.v19i4.4944>
- Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new frequency measure for Dutch words based on film subtitles. *Behavior Research Methods*, 42, 643–650.
- Körner, A., Bakhtiari, G., & Topolinski, S. (2019). Training articulation sequences: A first systematic modulation of the articulatory in-out effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45, 1725–1732.
- Körner, A., & Rummer, R. (2022). What is preferred in the in–out effect: Articulation locations or articulation movement direction?. *Cognition and Emotion*, 36, 1–10.
- Levelt, W. J., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition*, 50, 239–269.
- Lindau, B., & Topolinski, S. (2018). The articulatory in-out effect resists oral motor interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44, 209–220.
- Lüdtke, J., & Jacobs, A. M. (2015). The emotion potential of simple sentences: Additive or interactive effects of nouns and adjectives?. *Frontiers in Psychology*, 6, 1137.
- Lupyan, G., & Winter, B. (2018). Language is more abstract than you think, or, why aren't languages more iconic? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373, 20170137.
- Maschmann, I. T., Körner, A., Boecker, L., & Topolinski, S. (2020). Front in the mouth, front in the word: The driving mechanisms of the in-out effect. *Journal of Personality and Social Psychology*, 119, 792–807.
- Mohammad, S. (2018, July). Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 English words. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 174–184). Stroudsburg, PA: Association for Computational Linguistics.
- Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., Van Schie, K., Van Harmelen, A. L., ... & Brysbaert, M. (2013). Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behavior Research Methods*, 45, 169–177.
- Perniss, P., Thompson, R., & Vigliocco, G. (2010). Iconicity as a general property of language: Evidence from spoken and signed languages. *Frontiers in Psychology*, 1, 227.
- Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, 8, 364–382.
- Rossi, P., Pantoja, F., Borges, A., & Werle, C. O. C. (2017). What a delicious name! Articulatory movement effects on food perception and consumption. *Journal of the Association for Consumer Research*, 2, 392–401.
- Rude, S., Gortner, E. M., & Pennebaker, J. (2004). Language use of depressed and depression-vulnerable college students. *Cognition & Emotion*, 18, 1121–1133.
- Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior*, 23, 383–404.
- Silva, R. R., & Topolinski, S. (2018). My username is IN! The influence of inward vs. outward wandering usernames on judgments of online seller trustworthiness. *Psychology and Marketing*, 35, 307–319.
- Thompson, R. L., Vinson, D. P., Woll, B., & Vigliocco, G. (2012). The road to language learning is iconic: Evidence from British Sign Language. *Psychological Science*, 23, 1443–1448.
- Topolinski, S., & Boecker, L. (2016a). Mouth-watering words: Articulatory inductions of eating-like mouth movements increase perceived food palatability. *Appetite*, 99, 112–120.
- Topolinski, S., & Boecker, L. (2016b). Minimal conditions of motor inductions of approach-avoidance states: The case of oral movements. *Journal of Experimental Psychology: General*, 145, 1589–1603.
- Topolinski, S., Maschmann, I. T., Pecher, D., & Winkielman, P. (2014). Oral approach-avoidance: Affective consequences of muscular articulation dynamics. *Journal of Personality and Social Psychology*, 106, 885–896.
- Topolinski, S., Zürn, M., & Schneider, I. K. (2015). What's in and what's out in branding? A novel articulation effect for brand names. *Frontiers in Psychology*, 6, 585.
- Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, 45, 1191–1207.
- Weintraub, W. (1989). *Verbal behavior in everyday life*. New York: Springer.

- Westermann, R., Spies, K., Stahl, G., & Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology*, 26, 557–580.
- Wichmann, S., Holman, E. W., & Brown, C. H. (2010). Sound symbolism in basic vocabulary. *Entropy*, 12(4), 844–858.
- Winter, B., Perlman, M., Perry, L. K., & Lupyan, G. (2017). Which words are most iconic?: Iconicity in English sensory words. *Interaction Studies*, 18, 443–464.
- Whissell, C. (1999). Phonosymbolism and the emotional nature of sounds: Evidence of the preferential use of particular phonemes in texts of differing emotional tone. *Perceptual and Motor Skills*, 89, 19–48.
- Yelp Inc. (2019). *Yelp open dataset*. Retrieved from <https://www.yelp.com/dataset>