

# Distributional profiling and the semantics of modifier classes

Martin Schäfer  
Universität Tübingen

## Abstract

The versatility of modifiers makes their analysis a challenge. Their interpretation depends on their on lexical semantics, their syntactic position, and the semantics of the material they combine with. This paper explores what distributional profiling reveals about four modifiers that are traditionally taken to correspond to different semantic classes. The profiles of four different adjectives are presented and the results are qualitatively interpreted. While the results show that assumptions made in the literature are partly reflected in the distributional data, they also show a host of detail that requires other explanations.

## 1 Introduction

Four English adjectives, *lucky*, *quick*, *slow*, and *wise*, were selected that fulfill the following syntactic and semantic criteria: they occur in attributive as well as predicative position. Further, in the predicative position, they allow the combination with the *to*-INF(initival), and they all have a corresponding *-ly* form. Semantically, they are intended to cover a span of different lexical properties, making them representative of corresponding semantic classes. Their *-ly* variants are discussed in the adverbial literature as occurring with distinct readings. Further, at least for *wise* and *lucky* the occurrence in the *to*-INF construction as been described as expressing the same meaning as their occurrence as *-ly* adverbs.

This paper focuses on six patterns, three adjectival and three adverb patterns. They are illustrated for *wise/ly* in (1) and (2), with examples from the British National Corpus.

- (1) a. attributive adjective [attrib]  
'A most wise precaution,' Karl said.  
[A7A 3043]

- b. *to*-infinitival [INF]  
It is always wise to look towards those who inspire and impress you.  
[C9Y 1918]
- c. predicative [pred]  
If that is so, he is wise. [AKY 830]
- (2) a. sentence initial adverb [advSI]  
Wisely, Bright has included biographical entries of dead linguists only.  
[J7K 33]
- b. preverbal adverb [advV]  
The CO wisely decided not to notice this particular instance of it.  
[ACE 2163]
- c. postverbal adverb [Vadv]  
We help you choose wisely.  
[A65 1983]

Note that they are all treated as mutually exclusive, e.g., pred means predicative but not followed by the *to*-INFINITIVAL. Section 2 presents a short overview of work on modifier classes as relevant to this investigation. Section 3 presents the distributional analysis. The results are discussed in Section 4, Section 5 concludes.

## 2 Modifier classes

Modifiers have been classified via syntactic, morphological and semantic criteria. The considerable interplay of these criteria is one of the main difficulties for any analysis. Starting at the morphological level, it is unclear whether the relationship between adjectival base form and the *-ly* form should be seen as one of derivation or inflection. On balance, an inflectional characterization seems preferable, cf. Giegerich (2012); Plag (2018) pace Payne et al. (2010). At a purely syntactic level, the four items of interest are not qualitatively distinguishable as they all occur in all six patterns (as opposed to distributionally restricted items, e.g. the *alive* class that does not occur at-

tributively). The interesting distinctions are semantic in nature. At the level of lexical semantics, at least *quick* is seen as an event predicate, cf. Bücking and Maienborn (2019). This classification is based on conceptual necessity within their overall analysis. A similar judgement should hold for its antonym *slow*, too. I am not aware of specific discussion of the lexical semantics of *wise* and *lucky*, nor of a distribution based classification of these items. The different adjectival usages are not distinguished in detail for *quick* and *slow*, but it has been pointed out that not all readings in the attributive position are available in the predicative position, cf. Larson (1998, 18); Bücking and Maienborn (2019, 35f). For the adverbs, a link between semantic interpretation and syntactic position has been assumed at least from Jackendoff (1972) onwards. For the *quickly* usages in (3), (29) in Travis (1987), Travis distinguishes between (2-a-b) and (2-c-d).

- (3) a. **Quickly** John will be arrested by the police.  
 b. John **quickly** will be arrested by the police.  
 c. John will be **quickly** arrested by the police.  
 d. John will be arrested **quickly** by the police.

In her words, in the first two positions, “*quickly* appears to be modifying the event of the arrest while in (29c,d) *quickly* modifies the process of the arrest. In other words, in (29a,b), the arrest will happen right away. In (29c, d), the manner of the arrest will be hurried.” (underlining changed to italics by me) (Travis, 1987, 11). This distinction seems not to be available in this form for *slowly*, though I am not aware of explicit discussion in the literature. The most comprehensive treatment of the link between adverb position and adverb interpretation in English is still Ernst (2002), where high and low readings are linked by a manner rule. Where available, sentence initial position is linked to a high reading, and postverbal position to a low reading. In its high reading, *wisely* and *luckily* are sentence adverbials, *wisely* subject or rather agent-oriented, *luckily* from the evaluational subclass of speech-act adverbials (cf. Maienborn and Schäfer (2011)). The semantic difference to manner usages is, among other things, revealed through paraphrases:

- (4) a. Wisely, Bright has included biographical entries of dead linguists only. [J7K 33]  
 b. It was wise of Bright that they included biographical entries of dead linguists only.  
 (5) a. Luckily the flies had gone by now. [AON 2400]  
 b. It was lucky that the flies had gone by now.

In contrast, if manner usages are available, paraphrase like ‘in an ADJ manner/the way in which .. BE ADJ’ are more apt:

- (6) a. The old lady nodded wisely: ‘I thought so ... scientists would have tried it out on rats first.’ [A57 17]  
 b. The old lady nodded in a wise manner/The way in which the old lady nodded was wise.

While these paraphrases already show that sentences with the adverbs corresponds to sentences with the adjectives, both *wisely* and *luckily* have been explicitly linked to the *to*-INF pattern. Oshima (2009) points to the following two sentences with *wisely*, his (1ab), as being “roughly synonymous” (Oshima, 2009, 364) (cf. there for earlier literature making this point):

- (7) a. Wisely, John left early.  
 b. John was wise to leave early.

The corresponding two sentences with *luckily* show the same pattern, cf. (8), (6) in Oshima (2009, 363).

- (8) a. Luckily, John passed the exam.  
 b. John was lucky to pass the exam.

To my knowledge, the corresponding *quick to*-INF and *slow to*-INF patterns have not been explicitly discussed.

As this short discussion has shown, the four items have been discussed in the literature with different foci and to different extents. However, it should be clear that there are important differences in that only *wise* and *lucky* can be used as sentence adverbials, one as subject oriented, the other as evaluative adverbial. *Quickly* and *slowly* can both be used as manner adverbials, but whether their usage as higher adverbials is similar or not is

not clear. These difference motivated the choice, hoping that they stand for different classes of modifiers. All in all, little is known about the distribution of modifiers in actual corpora, and one main goal of this paper is to advance a method to close in on the specifics of the different items.

### 3 Distributional analysis

#### 3.1 Method

The distributional profiles compare the similarities of the four adjectives in its base form and across 3 adjectival and 3 adverbial patterns, as well as the adverb base form. The three adjectival patterns are a) the adjective in the attributive position, b) the adjective in predicative position followed by a *to*-INFINITIVAL, and c) the adjective in predicative position not followed by a *to* INFINITIVAL. The three adverbial patterns are a) the *-ly* form in sentence-initial position, b) the *-ly* form immediately preceding a main verb, and c) the *-ly* form in post-verbal position. To compare the different usages of each adjective, I used distributional semantics. The main idea behind this approach is to represent words exclusively via their distribution (for a comprehensive introduction and overview, cf. Sahlgren (2006)). There are many different ways this can be done, I proceeded as follows:

1. I first collected cooccurrence counts for each adjective, distinguishing between the three adjectival and the three adverb patterns. To extract the cooccurrence counts, I used the ukWaC. This is 2 billion word web-crawled corpus, where the web-crawl has been restricted to .uk domains (Ferraresi et al., 2008). The ukWaC corpus is part of speech tagged.
2. The cooccurrence counts were collected for the top 10,000 content words. The words had to cooccur with the target expression in the same sentence (Reddy et al., 2011).
3. The resulting raw counts were set to ratio of probability of context word given the target word to overall probability of context word (again, this follows (Reddy et al., 2011)).
4. To compare the similarity of the transformed cooccurrence counts, I used the cosine similarity between the resulting vectors.

#### 3.2 Results

The resulting similarities are shown for each of the four adjectives in Tables 1 through 4. A cosine

similarity of 1 indicates perfect similarity (the vectors are identical). The closer the value gets to 1, the more similar 2 vectors are. A cosine of 0, corresponding to a 90 degree angle, indicates unrelated scores. The NAs in the last column of Table 4 result from the rarity of postverbal *luckily*; the corresponding similarities could not be meaningfully calculated.

pattern	INF	pred	advSI	advV	Vadv
attrib	0.16	0.27	0.07	0.37	0.31
INF		0.22	0.09	0.37	0.36
pred			0.12	0.41	0.39
advSI				0.21	0.16
advV					0.73

Table 1: Cosine similarities between *quick*-usages.

pattern	INF	pred	advSI	advV	Vadv
attrib	0.2	0.58	0.14	0.46	0.44
INF		0.25	0.07	0.2	0.19
pred			0.12	0.33	0.33
advSI				0.49	0.41
advV					0.77

Table 2: Cosine similarities between *slow*-usages.

pattern	INF	pred	advSI	advV	Vadv
attrib	0.23	0.55	0.03	0.17	0.18
INF		0.15	0.08	0.16	0.18
pred			0.01	0.18	0.16
advSI				0.15	0.04
advV					0.21

Table 3: Cosine similarities between *wise*-usages.

Looking at the patterns across the four items, the tables show the following:

1. The attributive usage across all four adjectives shows little similarity to the *to*-INF (ranging from 0.16 to 0.23); it is more similar to the other predicative usages, with medium values for *slow* and *wise* (0.58/0.55). It is hardly related (*slowly/luckily*) or unrelated *quickly/wisely* to the sentence initial adverb pattern, but is relatively similar to both preverbal and postverbal adverbials for *quickly/slowly* (0.37/0.31 and 0.46/0.44), getting only low values for *wisely* and showing no similarity with preverbal

pattern	INF	pred	advSI	advV	Vadv
attrib	0.21	0.33	0.19	0.05	NA
INF		0.41	0.36	0.32	NA
pred			0.42	0.14	NA
advSI				0.19	NA
advV					NA

Table 4: Cosine similarities between *lucky*-usages.

*luckily*. The attributive pattern is most similar to the predicative usage for *slow/wise/luck* (0.58/0.55/0.33), while for *quick* it is most similar to the preverbal adverb (0.37).

2. The *to*-INF pattern varies much in its similarity to the other usages across the four adjectives. With the exception of *lucky*, it is less similar to the predicative pattern than the attributive is to the predicative pattern. *Lucky* also stands out when comparing the similarity between the *to*-INF pattern and the sentence initial adverb: the cosine value is 0.36, as opposed to no or hardly any similarity for the other three items. The similarity between *to*-INF and the pre- and postverbal adverbs is at the same level for *quick/slow/wise*, with higher values for *quick* (0.37/0.36 vs. 0.2/0.19, and 0.16/0.18, respectively). For *lucky*, the similarity to the preverbal adverb is 0.32.
3. The similarity between the predicative usage and the adverbs is moderate for *quick/slow*, with low similarity to the sentence initial adverb and the same degree of higher similarity to both the pre- and postverbal pattern. *Wise* shows a similar pattern on a lower level, with *lucky* being the odd one out with more similarity to the sentence initial pattern.
4. The sentence initial pattern shows medium similarity to the preverbal and postverbal adverb for *slow*, and a similarity around 0.2 to the preverbal adverb for *quick/wise/lucky*, and 0.16 similarity to postverbal *quickly* and no similarity to postverbal *wisely*.
5. The preverbal and postverbal adverbs are highly similar for *quickly/slowly* (0.73/0.77), showing little similarity (0.21) for *wise*. As mentioned above, for *lucky* this contrast does not apply.

## 4 Discussion

This section cannot meaningfully discuss all the nuances of the distributional data across the four adverbs. Instead, I will highlight three points in the results that are of special interest in view of the discussions and classifications in the previous literature.

### 4.1 Categorical distinctions and effects of lexical meaning

Across the different usages, there is no categorical break in similarity between the base forms and the *-ly* forms. This is in line with the view that *-ly* forms are more inflectional than derivational, which would not lead one to expect true meaning differences between these forms. In particular, both *lucky* and *wise* are more similar within their adjectival base forms than to any adverb form, with only the similarity between *to*-INF and pred for *wise* and that between pred and the sentence initial adverb for *lucky* the exception. For *slow*, the *to*-INF pattern is less similar to the other adjectives pattern than those to their most similar adverb. In contrast, for *quick* all adjectival forms are more similar to one of the adverbs forms than to any other adjectival pattern. Hypothesizing that the immediately preverbal and the postverbal patterns typically correspond to event-related usages, this supports the assumption of *quick* as primarily being an event predicate.

### 4.2 The *to*-INF pattern and the adverbial usages

The literature describes the sentence initial adverb pattern and the *to*-INF as roughly for *wise* and *lucky*. This leads one to expect some distributional similarity between the two patterns, which is indeed present for *lucky* (0.36), but completely absent for *wise* (0.08), which is here patterning with *quick* and *slow* (0.09 and 0.07 respectively). Whence this contrast? Closer inspection reveals that the *to*-INF pattern is special in that, in contrast to all other three adjectives, there is a high proportion of instances of subject extraposition in the data, cf. (9).

- (9)
- a. It is wise to use protective boots and an overgirth. [BPB 326]
  - b. It's wise to adapt accordingly. [CEF 1794]

While the exact reason for the preponderance of this pattern for *wise* and the effect of excluding this type on the distributional analysis must await further investigation, it is clear that the corresponding sentences are not paraphraseable by sentences with sentence initial *wisely*. Another surprising finding in relation to the *to*-INF pattern are the similarity values of 0.37 and 0.36 for the preverbal and postverbal *quickly*. This shows that the *to*-INF pattern does not specifically target sentence adverb correspondences. Closer inspection of the data does reveal that there are clear trends for a) verbs of communication and b) the higher interpretation of *quickly*, where the adverb indicates that a short time elapses between the moment of speaking and the onset of the action described by the INFINITIVAL, cf. (10).

- (10) a. And he is quick to point out that it was a joint decision to make a serious bid. [G39 1207]  
 b. The Japanese were quick to promise similar sums. [AB6 1170]

It is perhaps the absence of this higher interpretation for *slowly* which leads to lower similarity between the corresponding constructions.

### 4.3 The pre- and postverbal adverb positions

The literature assumes, where available, pairs of high and low readings for the adverbs. Of the positions tested, the sentence initial position is reserved for high readings, and the postverbal position for low readings. Due to sparsity of in the data, I could not distinguish between preverbal occurrences of the adverbs before the auxiliary complex and immediately before the main verb and opted instead to just consider the preverbal pattern. This pattern is ambiguous between high and low readings when no auxiliaries are present. While this complicates the interpretation of the observed similarities, there are interesting differences across the four items. *Quickly* shows a clear distinction in similarity between the sentence initial pattern and the other adverb patterns on the one hand (0.21 and 0.16) and the pre- and postverbal patterns among themselves (0.73). In contrast, while *slowly* shows the same high similarity between pre- and postverbal pattern, the sentence initial adverb shows moderate similarity to both the pre- and postverbal usage, too (0.49 and 0.41). This might be due to the unavailability of a dedi-

cated and clear higher reading for *slowly*. The values for *wisely* are relatively low across the board. It is not entirely clear why this is the case, but it might be related to the fact that here the high and low readings are conceptually very different, with the high reading not related to any internal aspect of the event. This view is supported by looking at the 5 top collocates of the adverb in the respective positions: for preverbal *wisely*, it is *avoid, refrain, decide, opt, and foresee*, all of which seem to disallow a manner interpretation of *wisely* in all circumstances. This contrasts with the top five postverbal collocates, *use, choose, nod, spend, and invest*, which invite manner interpretations. For *luckily* a clear manner reading is hard to conceive, which explains it hardly occurring postverbally.

## 5 Conclusion

This paper has shown that distributional profiling of modifier classes can offer new insights into the specifics of modifiers, and with them their corresponding classes. After presenting the distributional analysis of *quick, slow, wise, and lucky* across three adjectival and three adverb patterns, three findings were highlighted:

1. There is no binary cut in similarity between *-ly* and base forms. Nevertheless, subpatterns might point to underlying lexical semantic effects, for example the preference of *quick* for event modification.
2. The comparison between the *-to*-INF pattern and the adverbial pattern revealed that only *lucky* behaved as expected, while *wise* showed little similarity to the sentence initial adverbs, perhaps due to the effect of the subject extraposition. Surprisingly, the values for *quick* show that the *to*-INF pattern is not in itself fixed on a corresponding sentence adverbial function of the adverb.
3. Of the four adverbs, only *quickly* showed a clear binary reading distinction, while the other three departed from this pattern for other, partly unknown reasons.

Overall, this paper hopes to have shown that the combination of distributional profiling of different modifier usages with qualitative analysis is a promising step forward in the analysis of the semantics of modification.

## References

- Sebastian Bücking and Claudia Maienborn. 2019. Coercion by modification. The adaptive capacities of event-sensitive adnominal modifiers. *Semantics and Pragmatics*. Forthcoming in 2019.
- Thomas Ernst. 2002. *The Syntax of Adjuncts*. Cambridge University Press, Cambridge.
- Adriano Ferraresi, Eros Zanchetta, Marco Baroni, and Silvia Bernardini. 2008. Introducing and evaluating ukWaC, a very large web-derived corpus of english. In *Proceedings of the WAC4 Workshop at LREC 2008*, Marrakech. ELRA.
- Heinz J. Giegerich. 2012. The morphology of -ly and the categorial status of ‘adverbs’ in english. *English language and linguistics*, 16(3):341–359.
- Ray Jackendoff. 1972. *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge, Massachusetts.
- Richard K. Larson. 1998. Events and modification in nominals. In D. Strolovitch and A. Lawson, editors, *Proceedings from Semantics and Linguistic Theory 8*, Ithaca, NY. Cornell University. Download 11.11.03 from <http://semlab5.sbs.sunysb.edu/~rlarson/larson-papers.html>.
- Claudia Maienborn and Martin Schäfer. 2011. Adverbials and adverbs. In *Semantics. An international handbook of natural language meaning.*, volume 2, pages 1390–1420. Mouton de Gruyter, Berlin.
- David Y. Oshima. 2009. Between being wise and acting wise: A hidden conditional in some constructions with propensity adjectives. *Journal of Linguistics*, 45(2):363–393.
- John Payne, Rodney Huddleston, and Geoffrey K. Pullum. 2010. The distribution and category status of adjectives and adverbs. *Word Structure*, 3(1):31–81.
- Ingo Plag. 2018. *Word-Formation in English*. Cambridge Textbooks in Linguistics. Cambridge University Press, 2 edition.
- Siva Reddy, Diana McCarthy, and Suresh Manandhar. 2011. An empirical study on compositionality in compound nouns. In *Proceedings of the 5th International Conference on Natural Language Processing*, pages 210–218, Chiang Mai, Thailand. AFNLP.
- Magnus Sahlgren. 2006. *The Word-Space Model: Using distributional analysis to represent syntagmatic and paradigmatic relations between words in high-dimensional vector spaces*. Ph.D. thesis, Department of Linguistics, Stockholm University, Stockholm.
- Lisa Travis. 1987. The syntax of adverbs. In Benjamin Shaer, editor, *McGill Working Papers in Linguistics: Special Issue on Comparative Germanic Syntax*, volume 6, pages 280–310.