LUND UNIVERSITY

How similar are semantic categories in closely related languages? A comparison of cutting and breaking in four Germanic languages

Majid, Asifa; Gullberg, Marianne; van Staden, Mirjam; Bowerman, Melissa

Published in:
Cognitive Linguistics

DOI:
10.1515/COG.2007.007

2007

Link to publication

Citation for published version (APA):
Majid, A., Gullberg, M., van Staden, M., & Bowerman, M. (2007). How similar are semantic categories in closely related languages? A comparison of cutting and breaking in four Germanic languages. *Cognitive Linguistics*, *18*(2), 179-194. https://doi.org/10.1515/COG.2007.007

Total number of authors:
4

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

How similar are semantic categories in closely related languages? A comparison of cutting and breaking in four Germanic languages*

ASIFA MAJID, MARIANNE GULLBERG, MIRIAM VAN STADEN, and MELISSA BOWERMAN

Abstract

Are the semantic categories of very closely related languages the same? We present a new methodology for addressing this question. Speakers of English, German, Dutch and Swedish described a set of video clips depicting cutting and breaking events. The verbs elicited were then subjected to cluster analysis, which groups scenes together based on similarity (determined by shared verbs). Using this technique, we find that there are surprising differences among the languages in the number of categories, their exact boundaries, and the relationship of the terms to one another—all of which is circumscribed by a common semantic space.

Keywords: *cut and break; separation events; Germanic languages; English; German; Dutch; Swedish; verb semantics, categorization; cluster analysis; semantic map.*

1. Introduction

How similar, or different, are semantic categories in closely related languages? A widely held belief is that they are very similar. Take *break* (English), *brechen* (German), *breken* (Dutch), and *bräcka* (Swedish)—surely these words mean the same thing? They are cognates after all—they come from languages that are genetically related, and they can be traced back to a single Proto-Indo-European root (**bhreg*-). Moreover, the words refer to simple, concrete events, the meanings of which—at first glance—seem less likely to change than, say, the meanings of words for abstract things. Not only are the languages related, but their speakers also share very similar cultures, again suggesting that the semantic categories will be more alike than if the speakers came from very different cultures. All of these facts seem to bolster the assumption that the semantic categories of related languages are very similar to one another.

Cognitive Linguistics 18–2 (2007), 179–194 0936–5907/07/0018–0179
DOI 10.1515/COG.2007.007 © Walter de Gruyter

Table 1. *Some cognates used in the domain of C&B in English, German, Dutch, and Swedish*[2]

English	German	Dutch	Swedish
break	brechen	breken	bräcka
brittle		broos, verbrijzelen	bryta
(knife)		knippen	knipa
?clip			klippa
	kaputt	kapot	kaputt
peel		pellen	
hack	hacken	hakken	hacka
hew		houwen	hugga
hammer	hämmern	hameren	hamra
crush		krozen (obsolete)	krossa
half	halbieren	halveren	halvera
shell	schälen	schillen	?skiva
shear, shard		scheuren	?skära
slice		?slijten	slita
slay	schlagen	slaan	slå
(Old English sniðan)	schneiden	snijden	snida
	(zer)rissen, reissen		
saw	sägen	zagen	såga
deal	teilen	delen	dela
cut			kåta, kutå
rive	?reiben	?wrijven	riva
	scheiben		skiva
tear	zehren	teren (obsolete)	
chop	?kappen	?kappen	?kappa

However, there is some recognition that the relationship is more complex. Consider the terms for cutting and breaking (C&B, hereafter)[1] in English, German, Dutch and Swedish listed in Table 1. These cognates are identified by finding form-meaning pairs across the languages. The forms will not be identical because there are regular sound changes between languages (for example, postvocalic /t/ in English corresponds to postvocalic /s/ in German). Similarly, the meanings need not be identical. For example, English *black* is in part cognate with Russian *belo*, which means 'white'. Forms related to *black* in Slavic, Latin, and Greek refer to 'shining, flashing, brightening, whitening'. The shift in meaning of the English term comes from its specific derivation from the Germanic verb **blakaz* 'to have blazed, to have burned'. As an adjective the term would have meant burned or charred, leading to the current meaning black (Jeffers and Lehiste 1982). As with form changes, then, it is well accepted that meanings can change over time: semantic categories are subject to both contraction and expansion in reference (e.g., Traugott and Dasher 2001).

In the acquisition literature (e.g., Hill 1982; Jiang 2000), as well as in the psycholinguistic literature (e.g., De Groot 1992; Dijkstra, Grainger and van Heuven 1999), it is recognized that cognates are not isomorphic in meaning. Learning a foreign language exposes one to perilous false friends, where a similar form can be mistakenly taken to imply a similar meaning. A German speaker would be perplexed to read *bellen* next to a doorbell in the Netherlands: *bellen* in German means 'to bark', while in Dutch it means 'to ring'.[3] Yet, despite the appreciation by some that meaning is subject to change, and that related languages can have different semantic categories, there are very few techniques available for quantifying similarity or difference in meaning.

In this article we present one such approach for the synchronic comparison of word meaning in different languages,[4] and illustrate its application through the comparison of English, German, Dutch and Swedish categories of C&B. We show that even though these languages are closely related, there are differences in the number of categories, their exact boundaries and the relationship of the terms to one another. Yet despite this variation, the differences are not completely arbitrary, but circumscribed by a common semantic space.

We begin with a series of videoclips used to elicit descriptions of C&B events (Bohnemeyer et al. 2001; see Majid et al., this issue for a full description of the clips). The videoclips provide an "extensional grid" on to which we can map individual forms so as to compare terms across languages and map the interrelations of forms within a language. This allows us to investigate the overall semantic organization of a domain, and then to compare how similar and different semantic categories are across the four Germanic languages under study. For instance, we can examine questions such as: Do *break/brechen/breken/bräcka* pick out the same class of events? What is the relationship of each of these words to the other words within and across the languages?[5]

Previous work on the categorization of C&B events across languages (see Majid et al., forthcoming; Majid et al., this issue) has shown that typologically, genetically and areally diverse languages agree to a large degree on the categorization of C&B events. In an overall analysis of 28 languages, including English, German, Dutch and Swedish, there was considerable agreement in event classification. For example, all the languages distinguished between events on the basis of how predictable the locus of separation was (e.g., English *cut* (more predictable) vs. *break* (less predictable)). Most of the languages also distinguished tearing events, as well as snapping from smashing events. These dimensions of classification are respected by widely different languages, suggesting per-

haps that the categories of C&B events in the Germanic languages may be very similar to one another.

But further analyses comparing the *overall* similarity of the languages to one another in the classification of C&B events revealed a different aspect of the data: within the semantic space defined by the common structure across languages, Dutch, German and Swedish are more similar to one another than to any other language (not surprisingly), but English is quite different from the other Germanic languages. This indicates that there may, indeed, be differences in the semantic categories between these four languages.

2. General procedure

The analyses in this article are based on descriptions of the C&B videoclips from Bohnemeyer et al. (2001), elicited from five speakers of each language. Interviewed one at a time in their native language, participants watched the clips in a fixed randomized order and described what the Agent did. Each session was audio- and video-recorded for later transcription.

We focus on encoding of the core set of C&B scenes (N = 43). Clips depicting opening, peeling, etc., and clips depicting agentless events are excluded for reasons of space. Further, we limit our discussion to the verbs used to describe the scenes. Particles such as *off* and *apart* are also important to the description of state-change events in these languages (Talmy 2000), but an analysis of their semantics lies beyond the scope of this article.

3. Consistency between speakers

We began by examining how consistent speakers' descriptions were in each language. The question underlying this analysis is whether speakers within each language agree on how these C&B events should be linguistically categorized.

Number of verbs alone is one possible index of how much consistency there is for a particular stimulus event—the more verbs there are, the less consistency there is. But this measure alone can be under-informative about the data. Table 2 lists the verbs produced by speakers of English, German, Dutch and Swedish (ordered by their frequency of occurrence). There is no significant difference across languages in the total number of verbs produced (English N = 27, Dutch N = 24, German N = 29, Swedish N = 22) so one might be tempted to conclude that there is nothing noteworthy to say about consistency. But this would be premature.

Table 2. *Verbs produced to target C&B clips, with frequency of occurrence*

English		*German*		*Dutch*		*Swedish*	
verb	fre-quency	verb	fre-quency	verb	fre-quency	verb	fre-quency
cut	61	schneiden	48	slaan	48	hugga	44
break	33	schlagen	33	snijden	45	slå	39
chop	30	brechen	22	hakken	40	skära	33
smash	12	trennen	20	breken	22	bryta	16
tear	11	hacken	19	knippen	14	klippa	15
karate-chop	10	reissen	18	scheuren	11	slita	14
slice	8	teilen	10	maken	8	dela	10
snap	7	machen	6	trekken	5	krossa	7
make	6	hämmern	5	zagen	5	klyva	7
rip	6	sägen	5	delen	3	riva	6
hack	5	hauen	4	pulveren	2	såga	5
saw	5	stechen	3	gruizelen	1	skiva	3
slash	3	eindreschen	2	kappen	1	knäcka	3
bodge	2	halbieren	2	klieven	1	hacka	3
pound	2	schnibbeln	2	knakken	1	spetta	2
pull	2	biegen	1	knappen	1	sticka	2
separate	2	durchlöchern	1	knikken	1	ta	1
hit	1	entzweien	1	mishandelen	1	snitta	1
pierce	1	fetzen	1	prikken	1	splittra	1
poke	1	kleinern	1	slachten	1	hamra	1
pulverize	1	knicken	1	spietsen	1	göra	1
punch	1	loechern	1	splijten	1	mosa	1
slam	1	manschen	1	stoten	1		
spear	1	pflöcken	1				
split	1	ritzen	1				
stab	1	spalten	1				
trim	1	stampfen	1				
		stören	1				
		verletzen	1				

Table 3. *Two possible distributions of responses*

	verb1	verb2	verb3	verb4	verb5	verb6	verb7	verb8	verb9	verb10
Sample 1	10	10	10	10	10	10	10	10	10	10
Sample 2	91	1	1	1	1	1	1	1	1	1

We can supplement the information about number of verbs by looking at the distribution of the responses. Say we asked 100 people to describe a videoclip: a clip described with ten different verbs can nevertheless have very different distributions. For example, Sample 1 in Table 3 shows a

situation where ten verbs were used equally often to describe an event. In Sample 2, in contrast, one verb was used by 91 people and the other nine verbs were used only once each. We want to be able to capture the fact that in the second case there is more consistency of response, even though there are just as many different verbs as in the first case.

To capture both the number and the distribution of verbs, we used Simpson's diversity index to measure consistency. The higher the value of D, the more consistency there is in the responses offered.[6] We measured D for each videoclip and for each language separately. This tells us how consistent speakers in the four languages were in describing the clips. We then calculated the mean D for each language, to assess the overall consistency for each language. Swedish speakers were the most consistent ($D = 0.7$), followed by Dutch ($D = 0.6$), and then German ($D = 0.4$) and English ($D = 0.4$).

4. Cluster analysis of verbs

To compare the extension of the terms in the different languages, we created a videoclip-by-verb matrix for each language separately (with clips as rows and verbs as columns). For each clip, if a particular verb was used by any speaker for that scene then a one was coded; otherwise a zero.[7] The resulting matrices were analyzed using cluster analysis, a statistical technique that groups together items so that between-group variation is maximized while within-group variation is minimized.[8] In this technique, each videoclip starts out as a separate cluster, and then at every step the clusters are merged to form larger clusters based on similarity. In our analysis, the calculation of similarity is based on the use of verbs across videoclips. To the extent that clips are described with the same verb(s) they are more similar to one another and are more likely to be in the same cluster. Clips that are never described by the same verb(s) will end up in separate clusters.

The advantage of cluster analysis over other multivariate statistical techniques, such as multidimensional scaling or correspondence analysis, is that the hierarchical structure among items can be seen. This is pertinent to our study, since there is reason to think that English verbs of material destruction have a hierarchical relationship, for example, that *slice* is a hyponym of *cut* and *snap* is a hyponym of *break*. We want to be able to capture this structure where it exists.

The results of the cluster analyses for the four languages are presented in Figures 1a–1d. These are dendograms, where each videoclip is depicted as a separate row and identified by the clip number and a brief

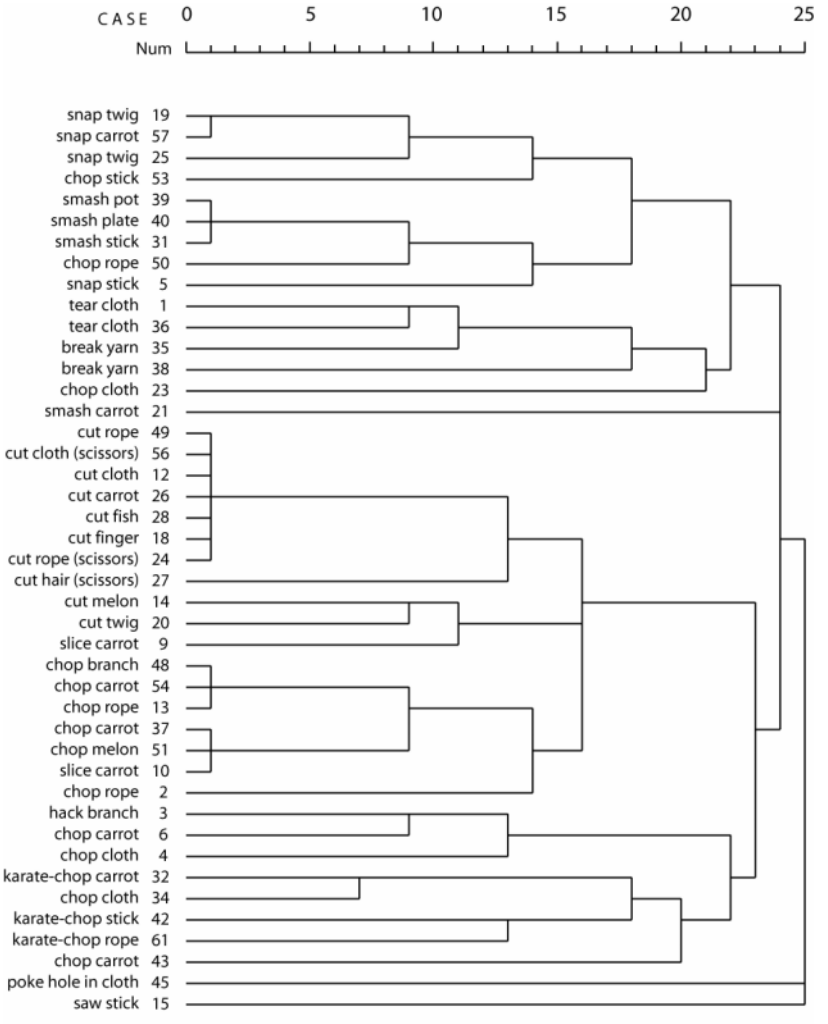

Figure 1a. *English dendrogram*

description.[9] The clusters capture the main groupings based on the distribution of verbs across the whole stimulus set. Videoclips that are the most similar to one another (because they were described by a single verb that was not used for any other clip) are clustered with the shortest leaves (lines) linking them from the left of the dendogram. Most clusters do not have very short leaves, since multiple verbs were used for the clips.

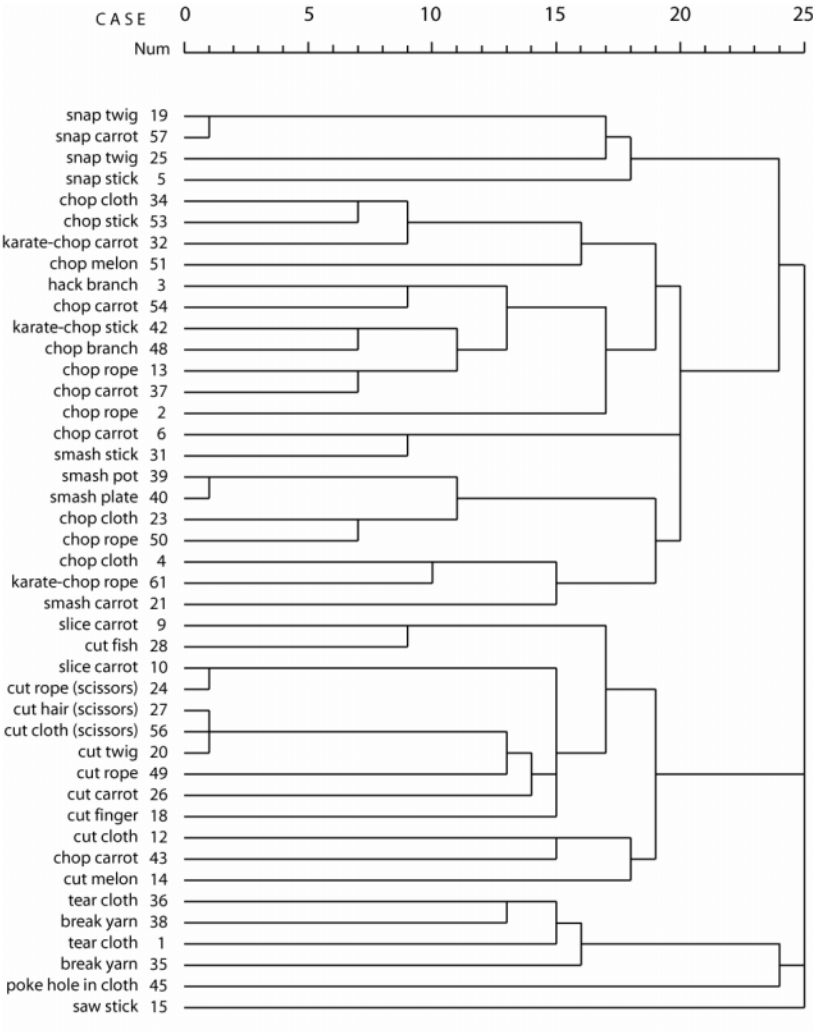

Figure 1b. *German dendrogram*

Some verbs were used only for a single clip, so they cannot be the basis for further grouping. If a cluster is embedded within a larger cluster, we can conclude that there is a hierarchical relationship amongst the verbs that were used to describe clips: there was at least one verb that was used for all the clips in the most encompassing cluster, and at least one verb used across the clips in the sub-cluster.

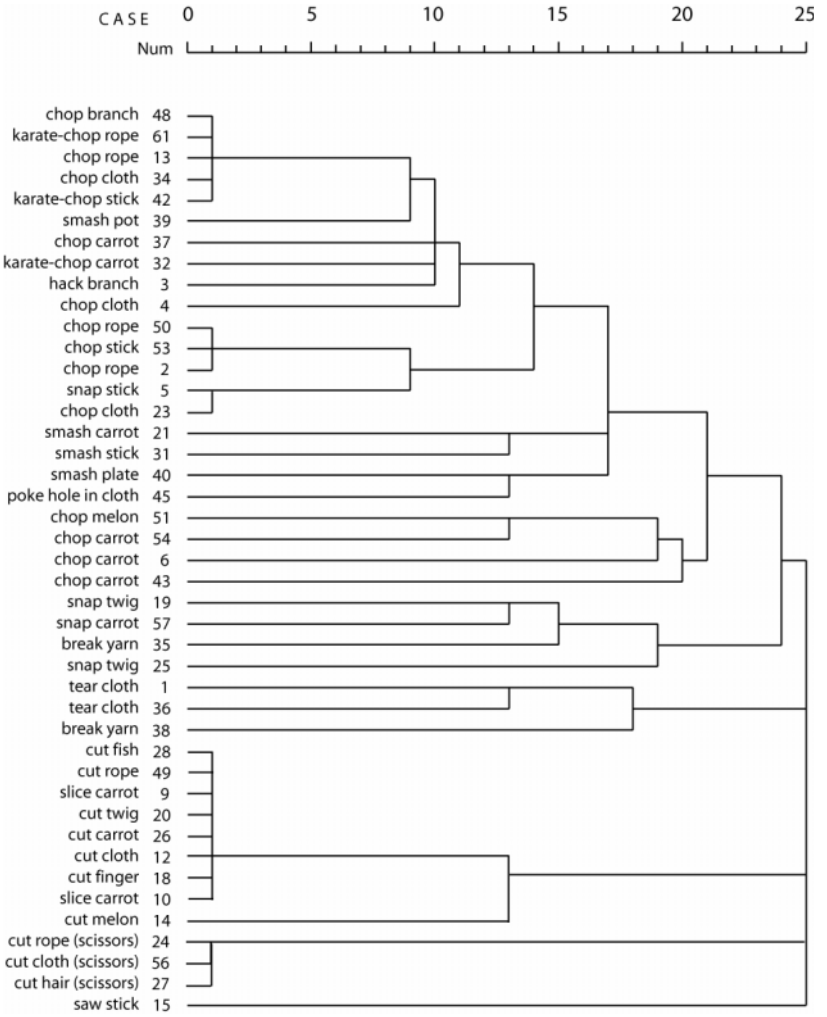

Figure 1c. *Dutch dendrogram*

For all four of the languages there is one clip which is completely on its own—this is known as the runt or entropy group. The runt item is clip 15, in which a man uses a saw to cut a branch in two. All of the speakers of each language used a single verb for that clip (*saw*—English, *sägen*—German, *zagen*—Dutch, *såga*—Swedish), and this verb was not used for any other clip in this stimulus set. This response pattern causes clip 15 to

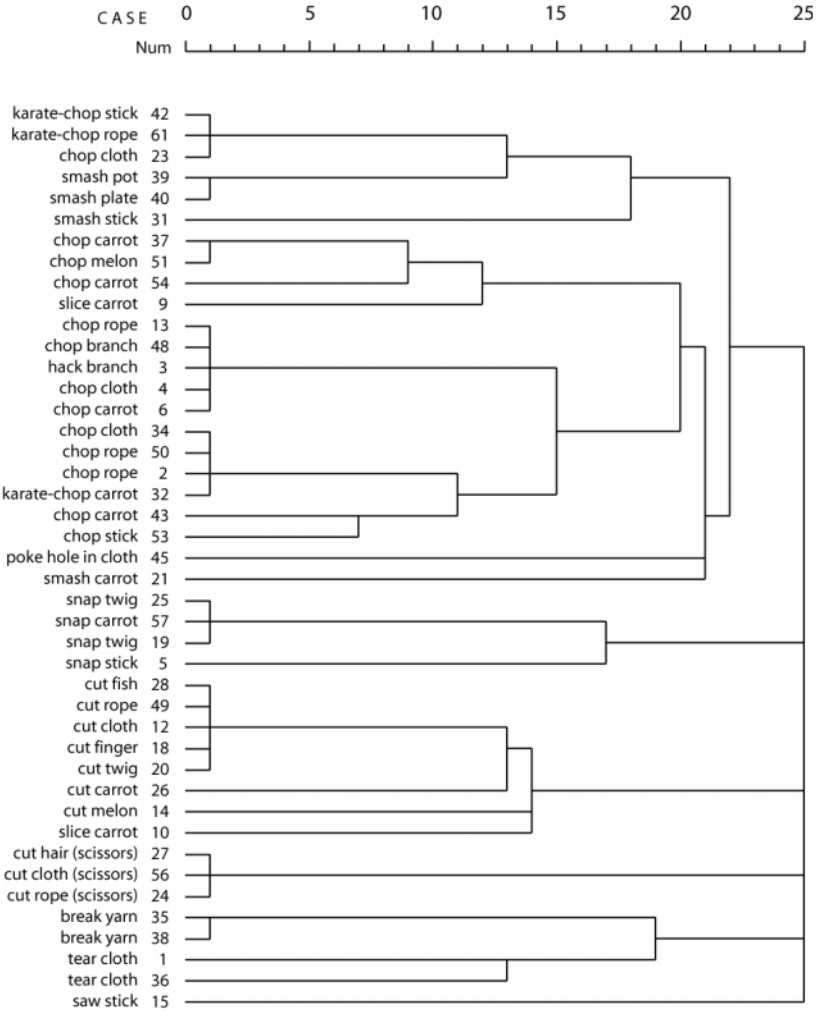

Figure 1d. *Swedish dendrogram*

be an outlier that forms a branch by itself; it is not grouped with the other clips until the very end of the procedure. This clip will not be discussed further.

4.1. *Descriptions of the individual languages*

4.1.1. *English.* There are two large clusters in the English dendogram (see Figure 1a). The top cluster is made up of breaking events, while the

bottom cluster is made up of cutting events. These were described by speakers as *break* and *cut*, respectively. Within these clusters there are sub-clusters. For the breaking cluster the sub-clusters consist of clips described with *snap*, *smash* or *tear*. Tearing forms part of a larger cluster of breaking events because there are some events (e.g., breaking/tearing a piece of yarn) which are called *break* by some speakers and *tear* by others, and this makes the overall similarity of the tear events higher to break events than to the cut events.

The cutting cluster includes all the separations involving knives or scissors. There is also a sub-cluster of chopping events (labeled, *chop*, *karate-chop* etc.). These are events in which a sharp instrument such as an axe is used to create a separation by means of a heavy blow. The chopping sub-cluster also includes events where a blunt instrument is used with a violent blow to create a clean separation, for example, a hammer, or a hand used to karate-chop.

4.1.2. *German.* The overall structure of German is different from that of English. Instead of two large clusters there are three (Figure 1b): a large breaking cluster, a cutting cluster, and a separate tearing cluster. The chopping cluster is found in the breaking cluster; recall in English it was subsumed by the cutting cluster. Further, while snapping events form a coherent sub-cluster, smashing events do not—instead, these are part of the larger cluster of chopping events. This is because German speakers largely restrict the verb *brechen*—the cognate of English *break*—to snapping events, whereas they sometimes describe smashing events as *schlagen* or *hacken*—verbs also used for some chopping events. The cutting cluster, which is associated with the verb *schneiden*, includes events in which an object is separated by a sharp instrument such as a knife or scissors. Finally, the tearing events form a separate cluster in which events of tearing cloth and breaking yarn are grouped together through the use of the verb *reissen*.

4.1.3. *Dutch.* Dutch has four distinct clusters (see Figure 1c). From top to bottom, the first cluster includes events of breaking, the second events of tearing, third events of cutting-with-a-single-blade, and the fourth events of cutting-with-scissors.

Tearing (*scheuren*), cutting-with-a-single-blade (*snijden*) and cutting-with-scissors (*knippen*) are self-explanatory. The breaking category is more like that of German than of English, in that it groups the chopping events with the snapping and smashing events rather than with the cutting events. Chopping clips were often described with the verb *hakken* (cognate with the English *hack*).

4.1.4. *Swedish.* Swedish has the highest number (five) of discrete clusters for categorizing C&B events: a large breaking cluster (*hugga*), snapping (*bryta*), cutting-with-a-single-blade (*skära*), cutting-with-scissors (klippa), and tearing (*slita*) (see Figure 1d). While Swedish, like Dutch and German, has a large breaking category that also encompasses chopping types of events, it differs from these languages in treating snapping events as a completely separate cluster.

4.2. *Comparison across languages*

The cluster analyses reveal the internal organization of the semantic field of C&B in each language. There are a number of points of difference. English has the deepest hierarchical structure of the languages. At the highest level there are only two major clusters. These are associated with two superordinate terms—*cut* and *break*. Within these two clusters there are a number of sub-clusters associated with subordinate terms such as *slice, chop, snap, smash*. Swedish, in contrast, has the flattest hierarchical structure, with five clusters, each of which is distinct. Notably, there are no superordinate terms equivalent to *cut* and *break* in English. While English speakers can choose to describe an event as either *cut* or *slice* for example, Swedish speakers do not have this option—choice of verb is determined by properties of the event, and the speaker has less room to construe the event in different ways.

This constraint in verb selection in Swedish is also reflected in consistency of responses. Recall that Swedish speakers were the most consistent in how they described the videoclips, while English speakers were much less so. We can now see why: the deeper hierarchical structure of the English C&B verb lexicon means that the same event can be described felicitously in alternative ways, while the flat structure of the Swedish lexicon rarely allows this.

The languages differ not only in the number of clusters they recognize, but also in how they group events together. Let us consider three examples. First, while English groups chopping events (i.e., events where a separation is caused by a sharp blow) with cutting events such as slicing and cutting-with-scissors, the other three languages group chopping events with breaking events such as smashing.

The second example involves tearing events. In German, Dutch and Swedish these form a higher-order cluster completely distinct from breaking events. But in English they are a subtype of breaking events because there is some overlap of terms used to describe tearing and breaking events.[10] Although tearing a piece of cloth is always described as *tear* and never as *break*, and snapping a twig is labeled *break* (or *snap*)

but never *tear*, some events—for example, separating a piece of yarn into pieces by hand—can be described by both verbs. Because of this overlap a higher-order cluster is found in English but not in the other languages.

Finally, the categories of cutting events also differ across the languages. Dutch and Swedish distinguish between cutting-with-scissors and cutting-with-a-single-blade. There is no superordinate term and consequently no higher-order cluster; the two types of events are treated as completely distinct. In contrast, English and German collapse over this distinction. English *cut* and German *schneiden* are indifferent to whether the blade is double, as with scissors, or single, as with a knife.

5. Discussion

There are a number of surprising differences in the semantic categories of C&B verbs in the Germanic languages. English, German, Dutch and Swedish vary in the number of categories they recognize, in the extension of these categories, and in the relationship of the terms to one another. Consider the cognate verbs we began with: *break* (English), *brechen* (German), *breken* (Dutch) and *bräcka* (Swedish). Our analysis shows a number of differences among these categories. English *break* picks out a superordinate category that subsumes finer distinctions such as those made by *snap* and *smash*. This verb is indifferent to how the effect was brought about, and it is also used to describe the destruction of a wide variety of objects, such as sticks, ropes, plates and yarn (see Pye 1994). *Brechen*, *breken* and *bräcka*, in contrast, all pick out a much more circumscribed set of events. None of these terms is a superordinate. German *brechen* and Dutch *breken* are used primarily for breaking long thin things by hand, i.e., snapping events. Swedish *bräcka*, on the other hand, is a rare verb used mainly for separating or cracking brittle, two-dimensional objects. Interestingly, it is never used to describe the scenes in the stimulus set in Bohnemeyer et al. (2001). The semantic category picked out by German *brechen* and Dutch *breken* exists in Swedish also, but it is not associated with the cognate term *bräcka*, but rather with an entirely different verb, *bryta*.

The differences reported in this article might seem at first glance to be at odds with the results reported by Majid et al. (this issue), who found that there was a core of common structure that languages share in their linguistic categories of C&B. Here we see that the categories vary quite substantially even in closely related languages. These findings can be reconciled by recognizing that the variation is indeed circumscribed by the common structure reported by Majid et al. According to Majid et al., lan-

guages make a distinction between events in which the locus of separation is predictable and those in which it is not. Predictability is a continuous dimension, with events involving separation with knives and scissors at one end, events of snapping and smashing at the other, and events in which a sharp blow causes the separation—chopping events—in the middle. In crosslinguistic perspective then, chopping events seem to be intermediate in the predictability of the location of separation.

This is reflected within the Germanic languages as well. English speakers group chopping events with cutting events involving knives and scissors, which are relatively high in predictability. Speakers of the other languages distinguish them from cutting events and group them together with breaking events such as smashing, which are relatively low in predictability. This crosslinguistic variability in the treatment of chopping events is also reflected in within-language speaker consistency. A median split was performed on the consistency data, so that clips that were consistently labeled were distinguished from clips that were diversely labeled. This analysis showed that consistent descriptions were given for events at the ends of the dimension discovered by Majid et al., while diverse labels were used for events that are not distinguished by the dimension. That is, even within a single language speakers were the most inconsistent in their descriptions of chopping events.

To conclude, the semantic categories of closely related languages can be very different from one another. Yet, at least in the domain of C&B events, this variation is played out within a common structure.

Received 24 June 2005
Revision received 8 December 2006

Max Planck Institute for Psycholinguistics, The Netherlands
University of Amsterdam, The Netherlands

Notes

* Any correspondence should be addressed to Asifa Majid, Max Planck Institute for Psycholinguistics, Postbus 310, Nijmegen, 6525XD, The Netherlands. We would like to thank the four reviewers and Barbara Malt for their insightful comments. Author's e-mail address ⟨Asifa.Majid@mpi.nl⟩.

1. We use "cutting" and "breaking" as short-hand descriptions for the real-world events that can be referred to by *cut* and *break*. By using these meta-descriptions for languages other than English, we do not mean to imply that the meanings are the same.
2. This list is based on various recent etymological dictionaries and studies on the history of (Indo-) Germanic languages. For English we consulted the *Oxford English Dictionary*. For Dutch the main sources are the *Etymologisch Woordenboek: De Herkomst*

van Onze Woorden and *Woordenboek der Nederlandsche Taal*. For German, *Indogermanisches etymologisches Wörterbuch* was used. For Swedish we used the *Svenska Akademiens Ordbok* (SAOB).

3. Thanks to Gertie Hoymann for this example.
4. We leave aside the fascinating issue of historical change due to space restrictions.
5. Note that in this approach we begin with an extensional space and investigate how it is divided, rather than beginning with the form and exploring all its possible meanings. These are two separate endeavors and should be viewed as complementary.
6. Simpson's diversity was calculated using the following formula: $D = (n_i - 1)/N(N - 1)$, where n_i is the total number occurrences of a particular verb (e.g., *cut*) and N is the total number of all verbs (i.e., *cut* plus *break* plus *chop*, etc.). In this study D is calculated separately for each language. D varies between 0 and 1.
7. A more graded measure, counting the number of speakers who gave that verb as a response, was not used, since there were so few participants.
8. Agglomerative hierarchical clustering was used with average linkage between groups (see Aldenderfer and Blashfield 1984 for an introduction to cluster analysis). The analysis was done using the Jaccard similarity ratio, according to which joint absence does not contribute to similarity (i.e., if two scenes are never described by the same verb, then those scenes are *not* considered to be similar to one another).
9. A fuller description of the videoclips can be found in Majid et al., this issue. For a copy of the video stimuli contact the first author.
10. Interestingly, children learning English apply *break* to events of tearing (Bowerman 2005).

References

Aldenderfer, M. S., and Roger K. Blashfield
1984 *Cluster Analysis*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07–044. Beverly Hills/London: Sage Publications.

Bohnemeyer, Jürgen, Melissa Bowerman and Penelope Brown
2001 Cut and break clips. In Levinson, Stephen C., and N. J. Enfield (eds.), *Field Manual 2001, Language and Cognition Group, Max Planck Institute for Psycholinguistics.* Nijmegen: MPI, 90–96.

Bowerman, Melissa
2005 Why can't you 'open' a nut or 'break' a cooked noodle? Learning covert object categories in action word meanings. In Gershkoff-Stowe, Lisa, and David Rakison (eds.), *Building Object Categories in Developmental Time: 32nd Carnegie Symposium on Cognition*. Mahwah, NJ: Lawrence Erlbaum, 33–62.

De Groot, Annette M. B.
1992 Bilingual lexical representation: A closer look at conceptual representations. In Frost, Ran and Leonard Katz (eds.), *Orthography, Phonology, Morphology, and Meaning*. Amsterdam: Elsevier, 389–412.

De Vries, Matthias, Lammert A. te Winkel et al.
1882–2001 *Woordenboek der Nederlandsche Taal*. Parts I–XXIX. Gravenhagen/Leiden.

Dijkstra, Ton, Jonathan Grainger, and Walter J. B. van Heuven
1999 Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language* 41, 496–518.

Hill, Robert J.
1982 *A Dictionary of False Friends*. London: Macmillan Press.
Jeffers, Robert J., and Ilse Lehiste
1982 *Principles and Methods for Historical Linguistics*. Cambridge, MA: MIT Press.
Jiang, Nan
2000 Lexical representation and development in a second language. *Applied Linguistics* 21, 47–77.
Majid, Asifa, James S. Boster, and Melissa Bowerman
forthc. The crosslinguistic categorization of everyday events: A study of cutting and breaking.
Majid, Asifa, Melissa Bowerman, Miriam van Staden, and James S. Boster
this issue The semantic categories of cutting and breaking events: A crosslinguistic perspective. *Cognitive Linguistics* 18(2), 133–152.
Pokorny, Julius
1994 *Indogermanisches etymologisches Wörterbuch*. Tübingen/Basel: A. Francke Verlag.
Pye, Clifton
1994 Breaking concepts: Constraining predicate argument structure. Unpublished manuscript. Department of Linguistics, University of Kansas.
Simpson, John A. and Edmund S. C. Weiner
1989 *The Oxford English Dictionary*. 2nd ed. Oxford: Clarendon Press.
Talmy, Leonard
2000 *Toward a Cognitive Semantics*. Cambridge, MA: MIT Press.
Traugott, Elizabeth C., and Richard B. Dasher
2001 *Regularity in Semantic Change*. Cambridge: Cambridge University Press.
Van Veen, Pieter A. F. and Nicoline van der Sijs
1997 *Etymologisch Woordenboek: De Herkomst van Onze Woorden*. 2nd rev. ed. Utrecht: Van Dale Lexicografie.