Peanut butter, compositionality and semantic transparency in loan translations

Eric Hoekstra and Frits van der Kuip

Abstract
This article presents an in-depth study of the Frisian loan *pindakaas* from Dutch *pindakaas*. This word can be literally glossed as ‘peanut cheese’, but it translates into English as ‘peanut butter’. The translation illustrates that the compound *pindakaas* is not compositional in Dutch and Frisian, that is, *pindakaas* is not a kind of *kaas* (‘cheese’). Because of its non-compositional nature, Dutch *pindakaas*, we argue, has not been borrowed into Frisian as *pindatsiis*, even though Dutch *kaas* in Frisian is *tsiis*. In contrast, compositional compounds featuring Dutch –*kaas* surface in Frisian with –*tsiis*, such as Dutch *schapenkaas*, Frisian *skieppe tsiis*. The non-compositional nature of *pindakaas* is shown to have a historical explanation. Independent evidence is cited from psycholinguistics supporting the claim that compositional compounds behave in a way that differs from non-compositional ones. Thus evidence is provided that borrowing is sensitive to compositionality in that elements of compounds are more easily left untranslated when their meaning is not predictable by compositionality from their usage elsewhere in the language.

Keywords
Compositionality, borrowing, loan blends, semantic transparency

1. Introduction
Some compounds and phrases obey semantic compositionality whereas others do not. For example, the compound *rice pudding* is compositional in that a rice pudding is a kind of pudding made of rice. The compound *black pudding* has two meanings, one of which is compositional whereas the other is not. It is compositional in case it means a pudding made of black ingredients. In that case, the meaning of ‘black pudding’ is predictable from knowing the meaning of ‘black’ and of ‘pudding’. In its non-compositional meaning, a black pudding refers to a ‘thick sausage made of the meat and blood of a pig’. In the latter case, the compound does not preserve compositionality since a black pudding in this sense is not a pudding. In this case, the idiomatic use of the compound causes loss of compositionality, seeing that the meaning ‘thick blood sausage’ cannot be predicted from knowing the meanings of ‘black’ and of ‘pudding’.

Compositionality may be defined as follows (Hoeksema, 2000, p. 851): “The meaning of a complex expression is determined by the meanings of its parts and the operations performed on those parts.” Hoeksema goes on to discuss the difference between sledge dog (which is a kind of dog) and dog sledge (which is a kind of sledge), pointing out that compounding requires an operation that distinguishes first elements of compounds from last elements. For the purpose of this article, we will only make use of a narrow interpretation of compositionality as involving a subset relation, as between sledge dog and dog. Natural
language does not always obey compositionality. As is well-known, compositionality gets lost in idioms and other kinds of conventionalised interpretations, such as in the example discussed above, *black pudding*. Furthermore, compositionality is not a binary principle but a gradual one: “There is a gliding scale from fully compositional (or fully motivated) to fully idiomatic (or conventional)” (Hoeksema, 2000, p. 855).

The purpose of this article is to show that this difference with respect to compositionality is relevant in case compounds or phrases are borrowed. Compositional elements of compounds and phrases are easily translated when they enter another language. However, a compound or phrase may be less than fully translated, containing an element which is recognisably stemming from a different language, a phenomenon also referred to as a loan blend (Haugen, 1950, p. 215; Haspelmath, 2009, p. 39). We will adopt the hypothesis that non-compositional elements have a tendency to be left untranslated when they enter another language. The present article deals with untranslated (parts of) compounds, more specifically with the word *pindakaas* (‘peanut butter’), which was borrowed into Frisian from Dutch. It will be argued that this word is not fully compositional, and that this explains why it was borrowed into Frisian without the compound element –*kaas* being translated to Frisian as –*tsiis*. In addition, we will also address the question why the Frisian purism for peanut butter, *nûtsjesmoar*, failed to catch on. In addition, we present other instances of non-compositional borrowings which are left untranslated in section 4.

Our analysis is within a usage-based framework (e.g. Geeraerts, Grondelaers and Bakema, 1994; Goldberg, 1995; Goldberg, 2006) in which linguistic factors may interact with cultural ones. In fact, it will be shown that the word *pindakaas* became non-compositional as a result of non-linguistic changes, i.e. changes in the outside world. We will also adopt the psycholinguistic notion of representational distance. Psycholinguistic experiments have provided strong evidence that bilingual speakers store words close together if they are similar in form and meaning (Dijkstra, 2003; Dijkstra, 2008; Smits et al., 2006; Smits et al., 2009). This applies not only to words from the same language but also to words from different languages. Thus, any two words, whether they belong to the same language or not, are stored close together or less close together, depending on the amount of similarity that exists between them.

Representational distance has both a formal dimension and a semantic dimension. For example, the Frisian word *wiis* (‘wise’) is stored close to its Dutch counterpart *wijs*, because they are very similar in form. Hence the formal distance between these two words is small. But these two words are also stored close together because they have the same meaning. Thus there is only a small semantic distance between these two words, as well. Similarly, there is only a small semantic distance between near synonyms like *sensible* and *wise*. In contrast, there is a large formal distance between *sensible* and *wise*, since the two words are very dissimilar in form. In this article, we will be concerned with semantic distance.

The notion of semantic distance helps us to understand what goes on with the two meanings of the example mentioned above, the *black pudding*. In its literal meaning, this expression is compositional. In this case, the meaning of *pudding* in *black pudding* is the same as the meaning of *pudding* in, for example, *warm pudding*; the semantic distance between these two instances of *pudding* will almost be zero. But now consider the meaning of *pudding* in the figurative meaning of *black pudding* (‘a thick blood sausage’). The latter instance of
pudding will be at a large semantic distance from compositionally used instances of pudding. Its literal or compositional meaning has receded to a greater or lesser extent into the background. Thus, when we say that a word is compositionally used, it means, from a psycholinguistic perspective, that the semantic distance between this use and other uses is practically zero. On the other hand, when we say that a word is non-compositionally used, it means that the semantic distance between this use and other uses is not fully predictable. It is not our purpose here to quantify the concept of semantic distance (but cf. Hay and Baayen, 2002; Heeringa, 2004; Versloot and Hoekstra, 2016). The notion of semantic distance helps us understand what it means psycholinguistically for an expression to be non-compositional.

2. The non-compositional case of kaas and pindakaas

2.1. The case of pindakaas

‘Nobody has been able to say anything deep about peanut butter,’ Prof. Jan Koster claimed in an interview (Koenen, 1990, p. 37; see also Koster, 1992). This claim is decidedly untrue in the field of word semantics, in which there are many interesting publications. We will see that the study of this word is interesting both from the perspective of word semantics (Donselaar, 2005) and from the combined perspective of word semantics, more specifically compositionality, and language contact, more specifically borrowing. With respect to lexical borrowing, the question arises why the Dutch word pindakaas (literally ‘peanut cheese’ > ‘peanut butter’) was borrowed into Frisian not as pindatsiis, but as pindakaas, whereas similar expressions with kaas (‘cheese’) are Frisianised to tsiis, such as Dutch Fricokaas (literally ‘Frico cheese’, where Frico is a brand name) or the loan blend cheddartsiis (‘cheddar cheese’). Tsiis is the Frisian equivalent of Dutch kaas, and it is in everyday usage. Frisian speakers practically always borrow from or via the dominant language Dutch. The case of pindakaas is a clear illustration of that. The word was not borrowed from English (peanut butter) or from German (Erdnusbutter). But why did Frisians borrow the entire Dutch word? Why didn’t they translate –kaas as –tsiis, which is, after all, a word generally known to speakers of Frisian? Is it just because peanut butter is in the Netherlands available only in jars labelled pindakaas? This seems unlikely, because Dutch labels like aardbeienjam (‘strawberry jam’) and rijstepap (‘rice pudding’) are translated in Frisian as ierdbeijam and rizenbrij. Or does the explanation relate to semantic properties of this particular word?

Disregarding Koster’s warning, let us first meditate on the word pindakaas itself. It is a compound consisting of the words pinda (‘peanut’$^3$) and kaas (‘cheese’). Applying compositionality to this word, we would expect that pindakaas is a kind of cheese. But it is not, as noted in passing by Sijens (2013). It lacks a prototypical property of cheese as it is known in The Netherlands. Cheese is prototypically firm and correspondingly, it can be sliced. $^3$ Peanut butter, in contrast, is prototypically soft and correspondingly, it can be spread. Seeing that pindakaas can be spread like butter, why didn’t Dutch language users directly coin the word pindaboter (‘peanut butter’)? The answer involves the legal protection of product names. By the time peanut butter was introduced on the market by the Dutch food companies, the Dutch word boter was legally protected, in that it was not allowed to use it for non-dairy products like margarine or peanut butter (Van Eerten et al., 2013). But why then choose for pindakaas? The answer involves the history of the concept underlying this word, as uncovered in Donselaar (2005), who writes from his personal experience:
Toen ik in 1958 voor het eerst in Suriname de pindakaas van daar proefde en verrast reageerde op de voortreffelijke kwaliteit in het bijzijn van een aldaar al langer wonende Nederlandse collega, zei die ‘Allicht, met die pittige smaak; en het is ook echt een kaas. Nederlandse pindakaas is gemalen poppenstrot’. De Surinaamse pindakaas van toen had een stevige consistentie. Je kon hem kopen bij een Chinese winkelier, die dan van een blokvormige massa een stuk van het gewenste gewicht afsneed. … Die pindakaas had een enigszins korrelige structuur en was niet smeerbaar.

(‘When I tasted pindakaas for the first time in Surinam and showed surprise at its excellent quality in the presence of a colleague who had lived there longer than me, he said: ‘Sure, it has a spicy taste; and it is really cheese. Dutch pindakaas is ground baby poo’. Surinam pindakaas was solid at that time. It could be bought at a Chinese grocer, who would slice the amount you wanted from the square mass. … The pindakaas had a somewhat granular structure and it could not be spread.’)

This excursion into cultural history makes it clear that the word obeyed compositionality when it was coined, and that events in the outside world were responsible for its non-compositionality. Pindakaas was originally a kind of kaas (‘cheese’) in the sense of being solid. It could be sliced and it could not be spread. It seems that the word pindakaas stayed on after its substance was turned into a soft spread by the food companies.

Thus Dutch pindakaas (‘peanut butter’) is different from, say, Dutch schapenkaas (‘sheep cheese’), or nagelkaas (‘clove cheese’), because both sheep cheese and clove cheese are a kind of cheese. If we consider the equivalents of these words in Frisian, this difference turns out to be relevant for borrowing:

Table 1. Some compounds with cheese in Dutch and Frisian

<table>
<thead>
<tr>
<th>Dutch</th>
<th>Frisian</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>schapenkaas</td>
<td>skieppetsiis</td>
<td>‘sheep cheese’</td>
</tr>
<tr>
<td>nagelkaas</td>
<td>nageltsiis</td>
<td>‘clove cheese’</td>
</tr>
<tr>
<td>boerenkaas</td>
<td>boeretsiis</td>
<td>‘farm cheese’</td>
</tr>
<tr>
<td>cheddarkaas</td>
<td>cheddartsiis</td>
<td>‘cheddar cheese’</td>
</tr>
<tr>
<td>schimmelkaas</td>
<td>skimmeltsiis</td>
<td>‘blue cheese’</td>
</tr>
<tr>
<td>pindakaas</td>
<td>pindakaas</td>
<td>‘peanut butter’</td>
</tr>
</tbody>
</table>

The Frisian Language Corpus gives 21 compounds with -tsiis, which are all equivalent to Dutch compounds with -kaas. There is only one type of compound with –kaas in Frisian: pindakaas, and this is exactly the one that is not compositional. The English translations make the same point: all compounds are translated with cheese, except the one that is not compositional since it doesn’t carry the meaning of ‘cheese’ for -kaas.
2.2. Failed alternatives for *pindakaas*: *pindatsiis* and **nútsjesmoar**

2.2.1. Failure of *pindatsiis*

There are 5 instances of *pindakaas* in the Frisian Language Corpus. There is, however, one instance of the fully frisianized form *pindatsiis*. Let us examine its context of use.

*Table 2. Only literary fragment containing the word *pindatsiis*, and its translation*

<table>
<thead>
<tr>
<th>Dan seit Marten:</th>
<th>Then Marten says:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Omke Gjalt, mei ik de <em>pinnekaas</em>?</td>
<td>Uncle Gjalt, can I have the <em>pinnekaas</em>?</td>
</tr>
<tr>
<td>Mar jonkje dochs! seit omke Gjalt.</td>
<td>O boy! says uncle Gjalt.</td>
</tr>
<tr>
<td>Marten wit, dat er it ferkeard dien hat.</td>
<td>Marten knows that he has made a mistake.</td>
</tr>
<tr>
<td>De <em>pindakaas</em>! seit er gau.</td>
<td>The <em>pindakaas</em>!, he corrects himselfs quickly.</td>
</tr>
<tr>
<td>Hy bedoelt: de <em>pindatsiis</em>, seit heit.</td>
<td>He means: the <em>pindatsiis</em>, father says.</td>
</tr>
<tr>
<td>Mar it is omke noch net nei ’t sin.</td>
<td>But this does not satisfy uncle Gjalt.</td>
</tr>
<tr>
<td>Jim moatte dy bern better Frysk leare, seit er.</td>
<td>You must teach the children to speak better Frisian, he says. <em>Pindakaas is nútsjesmoar.</em></td>
</tr>
<tr>
<td><em>Pindakaas is nútsjesmoar.</em></td>
<td>Can I have the <em>nútsjesmoar</em>, Marten asks.</td>
</tr>
<tr>
<td>Graach de <em>nútsjesmoar</em>, sei Marten dan.</td>
<td>Uncle Gjalt gives him the <em>nútsjesmoar</em>.</td>
</tr>
<tr>
<td>Omke jout him de <em>nútsjesmoar.</em></td>
<td>Suddenly Marten likes it much less.</td>
</tr>
<tr>
<td>Marten fynt it ynienen net sa lekker mear.</td>
<td></td>
</tr>
</tbody>
</table>

(from Eppie Dam, *Spinaazje mei spikers* (Boalsert: Kooperative Utjouwerij, 1986), p. 77.)

The quotation above comes from a children’s book. This example of *pindatsiis* is not quite to be taken seriously. This is clear from the context in which it is found, as illustrated by the quotation. It makes it clear that the normal word is *pindakaas*. The boy pronounces this word with Frisian phonology as *pinnekaas*. That is, he deletes the voiced plosive after a nasal, a characteristic rule of Frisian (both historically and synchronically, cf. Visser, 2000, pp. 167-168) and he reduces the vowel in the unstressed syllable to schwa, which is characteristic of Germanic languages in general (for Frisian: Visser, 2000, p. 150ff). This example makes it clear that *pindakaas* is the unmarked word for peanut butter, not *pindatsiis*.

In our analysis, the language user has difficulties forming *pindatsiis* from Dutch *pindakaas* exactly because the element -*kaas* is non-compositional. Thus non-compositionality is not only a property of the compound as a whole, but also of the part of it that is responsible for non-compositionality. This can be understood from a psycholinguistic perspective. The element -*kaas* in *pindakaas* has a meaning that is substantially different from -*kaas* as used in other compounds and phrases. Put differently, there is a substantial semantic distance between -*kaas* in *pindakaas* and -*kaas* as used in other compounds and phrases. In the compound *pindakaas*, it has a meaning close to ‘butter, spread’, and this meaning is not predictable from the meaning which *kaas* (‘cheese’) ordinarily carries. As a result, the
element -kaas in pindakaas is not well entrenched in the Frisian lexicon, not being semantically related to the concept kaas (‘cheese’) by compositionality.

To us as native speakers pindatsiis sounds unnatural, whereas we have no problem with skieppetsiis (‘sheepe cheese’), cheddartsiis (‘cheddar cheese’), Fricotsiis (‘Frico cheese’), and the like. The data in the Frisian Language Corpus resemble our judgements. There are no further hits in the Corpus with pindatsiis. This provides support for the hypothesis that parts of compositional compounds are easily translated into another language whereas parts of non-compositional compounds are more easily borrowed without being translated. This analysis provides insight into the observation that *pindatsiis does not occur, whereas other compounds with -tsiiis are easily found (cf. also section 3.1).

2.2.2. Failure of the purism nûtsjesmoar
The quote above introduced the purism nûtsjesmoar, which is bookish Frisian for ‘peanut butter’. It is a purism intended to replace the already existing Dutch loan pindakaas. This purism is prescribed in recent dictionaries of Frisian (Fr.D. 1984, 1985, 2008), which fail to mention the widely used pindakaas. Nûtsjesmoar, though hardly used, has been marked as the official standard Frisian word in the new wordlist in Taalweb (2015). Language purification is the reason for the exclusion of pindakaas from the Frisian dictionaries.

In the literature, scepticism exists about the long term effect of language purification (Theissen, 1975; Geeraerts and Grondelaers, 1997; Van der Sijs, 1999; Van der Kuip, 2010). Examples such as nûtsjesmoar, which is hardly ever used, raise the question how puristic a purism should or should not be (Van der Kuip et al., 2015). In other cases, there is rivalry between a Dutch-based loan and a native word, as is the case for pairs such as achter/efter (‘after, behind’), segen/seine (‘blessing’), doof/dôf (‘deaf’), koaning/kening (‘king’) (Sjölin, 1976; De Haan et al., 1978, especially p. 14, p. 30; on factors influencing success when two lexical items compete for the same concept, see Zenner, Speelman and Geeraerts, 2014). The following chat passage, taken from the internet, provides a nice illustration of the fact that the purism nûtsjesmoar is not well-known:
Table 3. Internet discussion on pindakaas / nútsjesmoar

<table>
<thead>
<tr>
<th>- Oké, maar wat is pindakaas nou int frysk??</th>
<th>- OK, but what is pindakaas in Frisian??</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Wy sizze gewoan pindakaas, dus ik wit it eins net.</td>
<td>- We just say pindakaas, so I actually don’t know.</td>
</tr>
<tr>
<td>- hieck tocht jir ut antwurd te heljen, nee</td>
<td>- Had I thought to get the answer here, but no</td>
</tr>
<tr>
<td>jir…niks… en ik wit wel sekur dat er een frysk wurd for is.</td>
<td>… nothing … and I know for sure that there is a Frisian word for it.</td>
</tr>
<tr>
<td>- Apenutensmots as sa miskien.</td>
<td>- Monkey nut sauce or so may be.</td>
</tr>
<tr>
<td>- Tadaaaa fûn mei de fertaalmachine hjirboppe it topic.</td>
<td>- Tadaaaa found with the translating machine above this topic.</td>
</tr>
<tr>
<td>It is&gt;&gt;&gt;&gt;&gt;&gt;&gt;&gt;&gt;&gt;&gt;&gt;&gt;&gt;nútsjesmoar.</td>
<td>It is&gt;&gt;&gt;&gt;&gt;&gt;&gt;&gt;&gt;&gt;&gt;&gt;&gt;&gt;nútsjesmoar.</td>
</tr>
</tbody>
</table>

(http://www.blauwefeniks.nl/forum/messages/211/8987.html)

This discussion underlines the fact that the word nútsjesmoar is rarely used in the spoken or written language. It occurs 11 times in the FLC; this may seem a lot, but all quotes come from just two sources, and its context of use is telling. The source having one instance of nútsjesmoar mentions that it is a difficult word. Of course, this source does not intend to say that the concept is difficult, since the concept of peanut butter is easy enough to grasp. It intends to express the fact that the word is by and large unknown to Frisian speakers. The other source, containing 10 instances, is the one from which the quoted passage in table 2 has been taken, and this source also made it clear that nútsjesmoar is not a well-known word, but an example of bookish Frisian.

Nútsjesmoar is not only a failure in actual practice, it is also a failure from a psycholinguistic point of view. The fact of the matter is that the word smoar (‘animal fat’) is obsolete in Frisian. Most speakers do not know it, and those who know it, hardly ever use it. Thus the meaning of nútsjesmoar (‘peanut butter’) is trivially not predictable from its constituent parts, hence it is not compositional. And even if the meaning of the constituent parts were known, the compound’s meaning could not be predicted from its constituent parts. The second element -smoar means ‘animal fat’, not ‘baking fat’ or ‘plant fat’, and, the first element doesn’t mean ‘peanut’, but simply ‘nut’. As a result, the constituent parts of the compound are not well entrenched in the language’s lexicon. The form smoar (‘animal fat’) is hardly known, in part because of the semantic distance between ‘animal fat’ and ‘butter, spread’, but also because the word is rather old-fashioned and not well-known to speakers of Frisian. As a result, the compound’s meaning is not easily learned or formed by language users.

2.3. Conclusion

Compositionality is a notion pertaining to the semantic combinatory possibilities of elements from the lexicon. If meaning remains unchanged in the process of composing compounds or
phrases, we speak of compositionality. As a result, the constituent parts of a language expression are compositional in case there is no distance between the representation of the meaning of the constituent parts when occurring in other expressions and the representation of the meaning of the constituent parts in the particular expression under examination. That is what guarantees predictability. If this view is correct, non-compositionality is something gradual (as is also claimed in Hoeksema, 2000): the more semantic distance there is between a constituent element within a particular expression as compared to its use in other expressions, the greater the semantic distance, the greater the non-compositionality.

3. Discussion

3.1. Analysis

The discussion above testifies to the importance of the distinction between a word’s compositional meaning and its non-compositional meaning. Thus, Dutch *pindakaas* does not easily become *pindatsiis* in Frisian because it is not some kind of cheese. In contrast, Dutch *Fricokaas* (‘Frico cheese’) or *schimmelkaas* (‘blue cheese’) just appear in Frisian as *Fricotsiis* and *skimmeltsiis*, because they are a kind of cheese. The following schemes make this explicit by providing a simplified picture of Dutch and Frisian forms linked to the semantic representation they share.

**SCHEME 1A. COMPOSITIONAL COMPOUND WITH -KAAS/-TSIIS**

<table>
<thead>
<tr>
<th>Frisian form: Fricotsiis</th>
<th>----&lt;-----</th>
<th>tsiis</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEMANTIC: FRICOKAAS ----&gt;----- KAAS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘Frico cheese’</td>
<td>‘cheese’</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dutch form: Fricokaas</td>
<td>----&lt;-----</td>
<td>kaas</td>
</tr>
</tbody>
</table>

**SCHEME 1B. NON-COMPOSITIONAL COMPOUND WITH -KAAS/-TSIIS**

<table>
<thead>
<tr>
<th>Frisian form: pindakaas</th>
<th>tsiis</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEMANTIC: PINDAKAAS ----X----- KAAS</td>
<td></td>
</tr>
<tr>
<td>‘peanut butter’</td>
<td>‘cheese’</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Dutch form: pindakaas</td>
<td>kaas</td>
</tr>
</tbody>
</table>
The semantic representation uses Dutch labels for reasons of exposition. Scheme 1A graphically illustrates that semantically, Fricokaas ‘Frico cheese’ is a kind of cheese, as indicated by the line containing an arrow (>). Scheme 1B shows that pindakaas ‘peanut butter’ is not some kind of cheese, as indicated by the line containing an X. This failure of compositionality is an obstacle to translating the Dutch compound element -kaas ‘cheese’ into Frisian -tsiis. Elements of a compound are left in the original language, in case they are not associated with the meaning which they ordinarily carry.

3.2. Why not apenutekaas?
Frisian has two words for peanut. There is the loan pinda, but there is also the native compound apenút (‘monkey nut’). Hence we predict that speakers familiar with the latter word should be able to use it in the translation of pindakaas. The -kaas element, of course, remains as it is, being non-compositional, but the element pinda- could be translated as apenute-, where the schwa is added as a linking suffix to left-hand members of compounds (Hoekstra, 1996). This should yield the following translation: apenutekaas. However, the word apenutekaas is found neither in the Frisian Language Corpus nor in Frisian texts on the internet. Frisian rarely features compounds containing as a left-hand member the form apenute-. The Frisian Language Corpus contains only one instance, the word apenutebakje (‘peanut bowl’). The dictionary of the Frisian language (WFT, 25 volumes) presents just 4 examples of compounds with the left-hand member apenute-:

(1) Four compounds with apenute- ‘peanut’:

- apenutedop ‘peanut shell’ 1 citation (1960)
- apenutekoekoek ‘peanut coockie’ 1 citation (1905)
- apenuteplant ‘peanut plant’ 1 citation (1938)
- apenútsjetsiis ‘peanut butter’ 1 citation (1946)

As a fifth example, a native speaker of Frisian reports the use of apenutebus (‘jar for peanuts’).

The non-occurrence of apenutekaas is backed up by native speaker intuitions. It is unclear why *apenutekaas should sound so awkward. It was observed that apenute- rarely occurs as a left-hand member of compounds. This begs the question why apenute- should be almost banned from compounds in the first place. There seems to be a relation between the non-occurrence of *apenutekaas and the general absence of compounds of this type, although we are unable to come up with a plausible explanation for this fact.

3.3. Luxury versus necessary loans and semantic drift
The literature on borrowing often makes a distinction between luxury loans and necessary loans (e.g. Onysko and Winter-Froemel, 2011; Haspelmath, 2009; Van der Sijs, 1996 and others). Necessary loans are words for which the borrowing language does not yet have a word. For example, the borrowing of the word computer into Dutch was a necessary loan because Dutch didn’t have a word with this meaning. Luxury loans are words for which the borrowing language does have an equivalent. For example, the borrowing of the word kids into Dutch is a luxury loan because Dutch already has the word kinderen (‘children’). Clearly, pindakaas is a necessary loan in Frisian, as there was no word for the concept of peanut butter.
in Frisian. Correspondingly, the word pindakaas is the unmarked form in Frisian, and it satisfies the criteria for unmarked forms (see Levinson, 2000, p. 137ff):

- It is relatively frequent in language use, that is, it is almost always used for the concept it expresses.
- It belongs to a neutral register.
- It carries no particular connotations.

This stands in stark contrast to the properties of the purism nútsjesmoar:

- It is not entrenched in language use, being rarely used for the concept it expresses.
- It does not belong to a neutral but a bookish register.
- It may carry unpleasant connotations, as in the example quoted above in table 2.

In case the marked form starts being used more often, it may develop a meaning or connotation different from the unmarked form. Bańko and Hebal-Jezierska (2014) investigated synonym pairs in Polish and Czech in their APPROVAL-project (APPROVAL stands for Adaptation, Psychological Perception and Reception of Verbal Loans). They relate differentiation in meaning and connotation to the fact that the form of the loan word is different, and is differently perceived by the language users. However, connotations do not so much depend on the form of the loan word as on the set of contexts in which it has been used, a point also mentioned by Bańko and Hebal-Jezierska (2014). The extensive Frisian example quoted in table 2 above shows that the context was responsible for the unpleasant connotation which the small boy attached to the purism nútsjesmoar. Successful luxury loans, analogously, tend to be used in modern contexts and this may affect their meaning. Put differently, the different perception of loan words can affect their semantic development. As a result, the loan word and the indigenous word can become near-synonyms, or even different words, as happened with the loan word strofa (‘stanza’, of a literary poem) and native word zwrotka (‘stanza’, of a folk or children’s song) in Polish. The loan word itself can also get a specific meaning (and different form), for instance the English loan people in French, which transformed to pipeul or pipole, meaning ‘famous people’ or even ‘famous person’ (Winter-Froemel, 2013). Finally, the English loan kids has a positive and fashionable connotation, by which the speaker profiles himself as a smart modern person, whereas the native Dutch kinderen is neutral in this respect. Compare the German words Kinder and Kids, where the latter ‘[i]s tied to contexts that portray the modern, emancipated child. This involves topics such as fashion, music and leisure time activities’ (Onysko and Winter-Froemel, 2011, p. 1562). It would be weird to use kids in a solemn context involving death of serious illness. Compare the following two Dutch sentences:

(2) We gaan met de kids op vakantie.
   (‘We go with the kids on vacation’.)
(3) # Een van hun kids heeft kanker.
   (‘One of their kids has cancer’.)

A sentence like (2) could be heard, but (3) sounds weird because the content of the sentence does not agree with the speaker’s profiling himself as a hip, modern sort of person. In a sense, (3) is in bad taste. In this way, the English loan kids is differentiated from the Dutch equivalent kinderen. Note that there might be speaker differences here, that is, there might be speakers for whom kids has become the unmarked form. The example makes it clear that a
borrowed word may have a different set of contexts of use in the recipient language, causing a semantic difference with the word in the original language. This happened in the case of peanut butter, as illustrated by the quotation in table 2 from the children’s book. There, the native word is originally a loan word, which is unmarked (pindakaas). The purism is marked (nútsjesmoar), and it picked up unpleasant connotations in a specific context of use.

4. Other instances of non-compositional loans

Compositionalty is a well-established semantic principle, but at the same time it is generally recognised that there are many expressions in natural language which violate compositionalty. This article provided an extensive discussion of the non-compositional loan pindakaas, but this is by no means the only example.

Frisian features the element -suer in the compound suerstof (‘oxygen’). The Dutch adjective zuur (‘sour’) is normally translated into Frisian as soer, so the expected form would have been *soerstof (stof ‘substance’ is the same word in Frisian and Dutch), but this form is infrequent in spoken Frisian. Note though that the noun zuurstof is not quite compositional since zuurstof (‘oxygen’) is not the same as zure stof (‘sour substance’). The following scheme illustrates the compositionalty failure of suerstof:

**Scheme 2. Non-compositional compound with -ZUUR/-SOER**

<table>
<thead>
<tr>
<th>Frisian form:</th>
<th>suerstof</th>
<th>soer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEMANTIC:</td>
<td>ZUURSTOF ----X----- ZUUR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>‘oxygen’</td>
<td>‘sour’</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dutch form:</td>
<td>zuurstof</td>
<td>zuur</td>
</tr>
</tbody>
</table>

Scheme 2 shows that zuurstof ‘oxygen’ is not zuur ‘sour’, as indicated by the line containing an X. This failure of compositionalty is an obstacle to translating the Dutch compound element zuur- ‘sour-’ into Frisian soer-. Dutch zuurstof is left untranslated and spelled in Frisian as suerstof, according to the rules of Frisian orthography.

Another example involves the element -loop in the Frisian loan overloop ‘landing at the top of the stairs’. Dutch over and lopen translate into Frisian as oer and rinne respectively, so the expected form would have been *oerrin. But this form is not found. The noun overloop is non-compositional since its meaning cannot be predicted from the meaning of its constituent parts: an overloop ‘landing’ is not a loop ‘run’.
A third example involves the constituent *leven in the phrase *syn leven* (*his life* > *‘ever*) in Middle Frisian. Middle Frisian consequently translates Dutch *leven* as *libben*, except in the expression *syn leven / libben*, where both forms can be found. This phrase is not compositional since it does not refer to any real life but instead it involves a quantification over time (see Hoekstra and Slofstra, 2013). Thus the word Dutch *leven* may be left untranslated in Frisian in case it is used non-compositionally.

### Scheme 4. Non-compositional Phrase with *leven/libben*

<table>
<thead>
<tr>
<th>Frisian form: syn leven</th>
<th>libben</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEMANTIC: ZIJN LEVEN --X--- LEVEN</td>
<td></td>
</tr>
<tr>
<td>‘ever’ ‘life’</td>
<td></td>
</tr>
</tbody>
</table>

Dutch features the compound *walrus* ‘whale horse’. English borrowed this word from Dutch as *walrus* (Durkin, 2014, p. 357). Again, no literal loan translation was coined, which would have been *whale horse*. The non-compositional interpretation of *walrus* provided an
obstacle to creating a loan translation. The word is non-compositional, as a *walrus* is not a horse. A similar example is the Dutch word *meerkat*. Literally, this is a ‘sea cat’, but its actual interpretation is ‘a kind of small monkey’. It was borrowed into English as *meerkat* (Durkin, 2014, p. 359). A *meerkat* is not a cat. Its non-compositional interpretation provided an obstacle to creating a loan translation.

Our analysis leads us to postulate that compositionality is a derived principle from a psycholinguistic point of view, that is, compositionality derives from the organization of our brain in terms of semantic distance. Furthermore, compositionality refers to a scale rather than to a binary opposition, again from a psycholinguistic point of view. A given element is more compositional if its meaning in a particular expression (a compound, a phrase) more strongly resembles the meaning it has in other expressions of the language.

This is completely in line with the fact that psycholinguistic tests have shown that non-compositional compounds react less well to cuing of their constituent parts than compositional compounds (Marelli, Amenta, Morone and Crepaldi, 2013). Furthermore, Hay and Baayen (2002) argue that non-compositional words tend to be accessed by whole word access (example: *listless*), whereas others are compositional, hence highly decomposable (example: *tasteless*). Decomposition arises as a result of parsing in perception. Parsing in perception depends on the independent existence of the two elements making up a compositional compound, and more specifically, these two elements must maintain the semantic interpretation which they have as independent words. Now, there is no independent word *list* meaning ‘sense of purpose’, and as a result, *listless* is not compositional. Correspondingly, it tends to be accessed by whole word access only in perception. Thus the perspective on compositionality taken here is independently supported by various psycholinguistic analyses.

Borrowing is sensitive to compositionality in the examples discussed above in that elements of compounds and phrases are more easily left untranslated when their meaning is not predictable by compositionality from their usage elsewhere in the language. The relation between different usages may be referred to as analogy (on analogy, see Hofstadter and Sander, 2013). Semantic analogies are easily recognised and used productively, but non-compositional constituent parts of the compounds or phrases are not easily related to translation equivalents in another language. Compositional words and phrases are predictable from their component parts, without any semantic distance which needs to be crossed. Such words and phrases are easily translated. However, if the meaning of a word or phrase requires an unpredictable jump from the meaning of the component parts, then there is a greater probability that the component parts responsible for non-compositionality are left in the original language. In such cases, the semantic distance is too great to be crossed. Thus the self-organisation of the brain is apparently sensitive to the compositional or non-compositional nature of the elements of compounds and phrases which are borrowed from one language into another.

**Acknowledgements**
We would like to express our gratitude to Siebren Dyk, Pieter Duijff, Jarich Hoekstra, Gertjan Postma, Hindrik Sijens and Nicoline van der Sijs for comments or discussion, and especially
to Ton van der Wouden and Willem Visser and two anonymous reviewers of Taal en Tongval for their extensive written comments.

References

Dam, Eppie, Spinaazje mei spikers (Boalsert: Kooperative Utjouwerij, 1986).


Dijkstra, Ton, Met andere woorden: over taal en meertaligheid (Nijmegen: Radboud Universiteit, 2008).


Theissen, Siegfried, De germanismen in de moderne Nederlandse woordenschat (Brussel: Belgisch Interuniversitair Centrum voor Neerlandistiek, 1975).


Van der Kuip, Frits, Pieter Duijff and Hindrik Sijens, ‘How puristic should a purism be?’, paper presented at the conference The Role of Lexicography in Standardization and Purification of Lesser Used Languages (Ljouwert/Leeuwarden: Fryske Akademy, 17 April 2015).

(to appear in a special of International Journal of Lexicography 2018)


WFT = Wurdboek fan de Fryske taal / Woordenboek der Friese taal (Ljouwert / Leeuwarden: Fryske Akademy, 1984-2011).


WNT = Woordenboek der Nederlandsche Taal


Information about the authors
Eric Hoekstra is linguist at the Fryske Akademy, KNAW / Royal Netherlands Academy of Arts and Sciences, Leeuwarden.
E-mail: ehoekstra@fryske-akademy.nl

Frits van der Kuip is lexicographer at the Fryske Akademy, KNAW / Royal Netherlands Academy of Arts and Sciences, Leeuwarden.
E-mail: fvanderkuip@fryske-akademy.nl

---

1 A related phenomenon, which we will not investigate here, involves folk etymology. Opaque compounds may be borrowed and changed so as to appear native-like in form, for instance Dutch *luipaard* ('leopard') from Latin *leopardus* (Van der Sijs, 1996, p. 492ff) or Frisian *titeloaso* ('daffodil') from older Dutch *tijdeloas, titelooas* (WFT s.v. titeloas, WNT s.v. tijlloas).

2 Van der Sijs (1996, p. 437) notes that the Dutch word *pinda* is a loan from Papiaments, which in turn derives it from Central African languages.

3 Speakers of Dutch and Frisian associate *kaas/tsiis* first of all with firmness. However, there are a few exceptions: Belgian Dutch *plattekaas* ('cottage cheese', ‘cream cheese’) (Van Dale), Dutch *smeerkaas, Frisian smartsiis* ('cheese to be spread > cream cheese’). The latter word has been claimed to be a loan translation from German *Schmierkäse* (Van der Sijs, 2002, p. 265).


5 Peanut butter is found in English alongside peanut spread. German features *Erdnussbutter* alongside *Erdnuss Creme*.

6 It can be found in digital form here: [http://gtb.inl.nl](http://gtb.inl.nl).

7 This compound is formed with the diminutive of *apenút* ('monkey nut > peanut’). It is not unusual for diminutives to be used in certain restricted types of Frisian compounds.