COGNITIVE STUDIES | ÉTUDES COGNITIVES, 18
Warsaw 2018

Article No.: 1717
DOI: 10.11649/cs.1717

Citation: Kafe, E. (2018). Persistent semantic identity in WordNet. *Cognitive Studies | Études cognitives*, *2018*(18). https://doi.org/10.11649/cs.1717

ERIC KAFE

MegaDoc, Denmark

kafe@megadoc.net

PERSISTENT SEMANTIC IDENTITY IN WORDNET

Abstract

Although rarely studied, the persistence of semantic identity in the WordNet lexical database is crucial for the interoperability of all the resources that use WordNet data. The present study investigates the stability of the two primary entities of the WordNet database (the word senses and the synonym sets), by following their respective identifiers (the sense keys and the synset offsets) across all the versions released between 1995 and 2012, while also considering *drifts* of identical definitions and semantic relations. Contrary to expectations, 94.4% of the WordNet 1.5 synsets still persisted in the latest 2012 version, compared to only 89.1% of the corresponding sense keys. Meanwhile, the splits and merges between synonym sets remained few and simple. These results are presented in tables that allow to estimate the lexicographic effort needed for updating WordNet-based resources to newer WordNet versions. We discuss the specific challenges faced by both the dominant synset-based mapping paradigm (a moderate amount of split synsets), and the recommended sense key-based approach (very few identity violations), and conclude that stable synset identifiers are viable, but need to be complemented by stable sense keys in order to adequately handle the split synonym sets.
Keywords: wordnets; semantic identifiers; sense keys; key violations; synsets; mappings

1 Introduction

1.1 Sense keys and synset offsets

Wordnets cover an increasing number of languages, and interoperate by using identifiers from the Princeton WordNet (PWN) lexical database (Fellbaum, 1998). PWN represents words by their generic form (*lemma*), grouping words that share the same meaning in synonym sets (*synsets*), so that different senses of the same lemma belong to different synsets. Thus, in PWN, *sense* is equivalent to synset membership, which constitutes a mapping between words and the *meaning* of the unique synonym set that each particular *word-sense pair* belongs to.

While the identifier for each synonym set, the *synset offset* (WordNet-team, 2010, Wndb), changes between each version of the database, each individual word sense has a stable identifier (the *sense key*) which, in principle, does not change across different PWN versions. So, according to the WordNet manual, "A *sense key* is the best way to represent a sense in semantic tagging or other systems that refer to WordNet senses" (WordNet-team, 2010, Senseidx).

In the Entity-Relation Model (Chen, 1976), each database key needs to be a unique attribute, which allows the retrieval of a single database entity. Before WordNet 1.5SC (1995), a few sense keys were not *unique* (and thus not keys), but in all later versions each sense key denotes only one sense of each word. Each word sense is a member of one and only one synonym set, so each sense key maps to only one synset offset in a given WordNet version. Additionally, each synonym set contains one and only one sense of each word that shares this sense, i. e. each synset offset corresponds to only one sense key of each word.

1.2 Mappings and updates

Primary keys are used as external keys by foreign databases, so they need to denote the same object over time, to avoid violating the **referential integrity** of external links. However, instead of using the recommended *sense keys* as their *foreign key*, almost all semantic databases that use PWN data are still linked to PWN through the ever-changing *synset offsets*, and thus bound to one particular version of PWN. This choice of foreign database key makes upgrades of the WN links difficult, and hinders the interoperability between resources that are bound to different PWN versions, such as many large scale ontologies (Niles & Pease, 2003; Suchanek, Kasneci, & Weikum, 2008) semantic knowledge bases (Navigli & Ponzetto, 2010), and most foreign language wordnets (Bond et al., 2014).

On the other hand, the *referential integrity* of the recommended sense keys has never been investigated. Kafe (2012) noted that the default meaning, i.e. the *lex_id* 0 (WordNet-team, 2010, Senseidx), of the word "C" had changed from an "alphabetic character" to a "programming language", between PWN 1.6 and 3.0. This violation of *sense identity* was discovered by comparing the output of a *sense key*-based mapping (which produced a false target), with the mappings described in Daudé, Padró, and Rigau (2001), which ignored sense keys, and produced the correct mapping target. So far, no other such error has been reported, and since *identity violations* are not allowed in theory, their practical extent could be negligible. But this assumption lacks confirmation.

Also, updating foreign language wordnets to a newer version of PWN requires additional lexicographic efforts, because the changes (splits, merges, deletions) in the PWN synsets do not always correspond to the composition of the foreign language synonym sets. So, in order to improve the precision of the mappings when updating between PWN versions, foreign language lexicographers need an accurate picture of the changes that occurred between these versions. But previous analyses have been limited to one PWN source and target pair: WN 1.5-1.6 (Daudé et al., 2001), WN 1.6-3.0 (Gonzalez-Agirre, Laparra, & Rigau, 2012), WN 3.0-3.1 (Vossen, Bond, & McCrae, 2016).

1.3 The stability of WordNet identifiers

The present study investigates the stability of the two essential entities of the PWN databases (the word senses and the synonym sets), by following their respective identifiers (the sense keys and the synset offsets) across all modern versions, ranging from WordNet 1.5 to the latest WordNet 3.1.1 for SQL.[1]

When the sense keys are unique and persistent, they permit to observe their groupings in synonym sets across PWN versions, and to trace how these synsets evolve in the database over time. Even though synset offsets change between versions, we can follow the sense keys of their members, and obtain a precise recension of all the splits, merges, additions and deletions that occurred between PWN versions, and thus estimate the lexicographic effort needed in order to achieve linguistically satisfying mappings.

[1] Version name suggested by Randee Tengi from the PWN team (personal communication).

2 Methods

2.1 Sense key identity

The primary input to our analysis is a PWN Sense Key Index (SKI) file, built from the *index.sense* files included in every PWN version since 1.5 (Kafe, 2017b, ski-pwn-flat.tab). In this form, the SKI is a complete table of tab-separated quadruples (sense key, WordNet version, part of speech, synset offset), linking every sense key to its synset offset in all PWN versions between 1.5 and 3.1.1.

By definition, if a sense key *k1* is a persistent database key, then the instances of *k1* present in different PWN versions *v1* and *v2* are the same key.

Definition: *Persistent sense key identity*

$$Key_{k1}^{v1} = Key_{k1}^{v2} \tag{1}$$

If these keys are the same, then they bi-imply each other, so a bidirectional mapping link exists between the two instances of *k1* present in the source synset *s1* and the target synset *s2*, which thus have a non-void *persistent* intersection. In theory, this inference is always valid, so mappings that only use this rule should not produce *false positives*.

Mapping Rule: *Sense key identity*

$$Key_{k1\in s1}^{v1} = Key_{k1\in s2}^{v2} \implies Key_{k1\in s1}^{v1} \leftrightarrow Key_{k1\in s2}^{v2} \tag{2}$$

2.1.1 Sense key identity violations

However, we saw previously that the default sense of the word "C" was different in WN 1.6 and 1.7, which shows that exceptions occur in practice, where two instances of the same sense key denote distinct senses of the same word across different WN versions.

Example of a sense key inequality:

$$Key_{c\%1:10:00}^{1.6} \neq Key_{c\%1:10:00}^{1.7} \tag{3}$$

This example constitutes a clear *identity violation*, and contradicts the definition of a persistent database key, so such keys do not satisfy the *referential identity* required by Rule 2. A *sense key identity violation* can thus be defined as the general formulation of the Sense Key Inequality above.

Definition: *Sense key identity violation*

$$Key_{k1}^{v1} \neq Key_{k1}^{v2} \tag{4}$$

We study these identity violations by considering drifts of identitical definitions in PWN updates:

Definitions	c%1:10:00	c%1:10:01
WN 1.6	**a general-purpose programing language closely associated with the unix operating system**	∄
WN 1.7	the 3rd letter of the roman alphabet	**a general-purpose programing language closely associated with the unix operating system**

In this example, the identity of the sense key C%1:10:00 is violated, because its definition changes so much that it does not define the same thing as previously. On the other hand, the identity of the new key C%1:10:01 is not violated, since this key did not exist previously, but the fact that its definition is identical to the source definition of the first key helps us discover the violation of that key.

We also study the same violation pattern in all the semantic relations included in the PWN data files, using a general two-step algorithm: find the set of *persistent sense keys* which are present in both the source and the target WN versions, and among the elements of this set, identify all *drifts* of *semantic essence* in the form of identical definitions or essential relations from a source key to any different target key of the same word, where the corresponding attribute of the source key is different in the target WN version. These two steps can also be applied in the reverse order, since the result is the intersection of two sets. But while there only exists one definition of the first set ($SenseKeys_{persist}$), several interpretations of $SenseKeys_{essencedrift}$ are possible. In this study, we define *essence drift* as the *drift* of any of the following WN attributes: *definitions, hypernyms, synonyms, antonyms, holonyms* and *pertainyms*, because these relations emerge as more *essential*, after initially considering all WN relations.

Finding sense key identity violations:

$$SenseKeys_{violated} = SenseKeys_{persist} \cap SenseKeys_{essencedrift} \tag{5}$$

The output is a set of probable key violations that can be examined manually. For example, in the WN 1.7 update, the identity of C%1:10:00 is also violated by its hypernyms.

Hypernyms	C%1:10:00	C%1:10:01
WN 1.6	**programing_language%1:10:00** **programming_language%1:10:00**	∄
WN 1.7	alphabetic_character%1:10:00 letter%1:10:01 letter_of_the_alphabet%1:10:00	**programing_language%1:10:00** **programming_language%1:10:00**

This confirms that a sense-specific mapping link is needed, in order to express that the WN 1.6 sense of C%1:10:00 is the same as the sense of C%1:10:01 in WN 1.7.

Mapping link:

$$Key^{1.6}_{c\%1:10:00} \leftrightarrow Key^{1.7}_{c\%1:10:01} \tag{6}$$

2.1.2 Sense and essence

To constitute an identity violation, a relation violation must violate the *essence* of a word sense, i.e. a necessary characteristic, which Aristotle distinguished from more contingent (or *accidental*) features. The latin translators of Aristotle used the term *essentia* to designate this central metaphysical concept, which has generated considerable controversy and diverging terminology since then (Cohen, 2016). Most notably, the scholastic philosopher John Duns Scotus opposed the essence of an individual (*haecceity*, or *haecceitas* in latin, literally "*thisness*"), and the related concept of *quiddity* (or *quidditas*, literally "*whatness*"), which is the essence common to a group (Cross, 2014). Although opposite according to Scotus, both terms are still synonyms in PWN 3.1, and hyponyms of {*kernel, substance, core, center, centre, essence, gist, heart, heart and soul, inwardness, marrow, meat, nub, pith, sum, nitty-gritty*}, and they share the same definition: *"the essence that makes something the kind of thing it is and makes it different from any other"*.

The notion of *essence* is necessary in order to identify violations of semantic identity, so these may remain elusive, until a precise and consensual definition of *essence* becomes available. Considering that this question already has remained open for almost twenty-five centuries, it seems safe to expect corresponding difficulties with the notion of *semantic identity*. In the meantime, the present study seeks to eschew the trap of a circular definition (violations of semantic identity are violations of essence and *vice versa*), by reviewing the violations of WordNet relations separately, in order to determine which ones violate the semantic identity of their arguments. As a practical example, even though the identity of the word "C" is violated, this does not affect the integrity of its hypernyms. The hypernyms express *what* "C" is, i. e. *quiddity*, which is a part of *essence*, while "C" does not add anything essential to the notion of "programming language". Thus, violations of superordinate relations like *hyperonymy* are more likely to be perceived as *essence* violations, since they express a necessary characteristic shared by all narrower word senses, while *hyponymy* violations only affect a part of the word sense which is not essential, since different hyponyms do not need to share that characteristic.

2.2 Analysis

2.2.1 The sense keys

After collapsing the *part of speech* and *synset offset* fields from the SKI database file into the 9-digit *synset id* format used in WNprolog (WordNet-team, 2010, Prologdb), we applied the built-in *xtabs* **cross-tabulatation** function in the *R* statistical environment (R-team, 2017), to obtain a table containing all the PWN versions as columns and all the sense keys as rows, with the *synset id* corresponding to each sense key and each PWN version in the cells, and *0* when the sense key was absent from the corresponding PWN version.

For each pair of consecutive PWN versions (see Table 2), we count the number of sense keys present in either the source version ($\mathbf{WN}_{source}$) or the target version ($\mathbf{WN}_{target}$), or both. Most sense keys *persist* in both versions, and their percentage expresses the *recall* of mappings that use only *Rule 2*. Sense keys that only appear in the source have been *removed* in the target, and those that only appear in the target have been *added* to the source. Violated sense keys are a special case: although present in both versions, they do not denote the same sense, and persist only in apparence. Here, we consider that their old sense is *removed* from the source WN version, while their new sense is *added* to the target WN, so that they are not counted as persistent.

The *persistent* and *removed* sense keys add up to $Total_{source}$, so we calculate their ratios as percentages of $Total_{source}$, which add up to 100. The *persistent* and *added* sense keys add up to $Total_{target}$, but their percentages do not add up to 100, because they are ratios of different totals. Both totals are identical to the *Word-Sense Pairs* reported in (WordNet-team, 2010, Wnstats).

2.2.2 Persistent, added and removed synonym sets

We analyse the evolution of the synonym sets, by considering whether their corresponding sense keys are present in either or both of the source and target PWN versions (see Table 3).

The source synset offsets of persistent sense keys have at least one translation in the target, and are counted as persistent synsets. Since violated sense keys are not considered persistent, they do not contribute to the persistent synsets. Source synset offsets that do not have a sense key present in the target correspond to *removed* synsets, while target synsets that do not have a sense key that was present in the source, have been *added* in the PWN update.

These figures and their percentages are calculated as for Table 2: the persistent and removed synsets add up to $Total_{source}$, and their percentages add up to 100. The synset totals are identical to those from each corresponding WN Stats manual page (WordNet-team, 2010, Wnstats). But, because of splits and merges, the number of persistent synsets in the source (i.e. the figure we use here) is not identical to the number in the target, which together with the number of *added* synsets, would add up to $Total_{target}$.

2.2.3 Split and merged synsets

In a mapping with unique pairs of (source, target) synset offsets, split synsets are those appearing more than once in the source column, while merged synsets are those appearing more than once in the target. The number of times that these synsets appear is a measure of the complexity of the split or merge operation. We indicate this size in Table 4 with a subscript, so that $split_2$ and $split_3$ are the number of synsets that were split in respectively two or three different target synsets. Similarly, $merged_2$ and $merged_3$ are the number of merges from two or three different source synsets. Some synonym sets are both split and merged, and we indicate their frequency as $\&^{merged}_{split}$.

The synonym sets counted as persistent here satisfy a minimal condition of stability, because they have at least one sense key present in both PWN versions. Extending Mapping Rule 2 allows to increase *recall*, by mapping removed sense keys to the target synset of their synonyms. The resulting Rule 7 generalizes the mapping previously established by Rule 2 for a single sense key *k1* from *s1*, to predict that its synonyms in version *v1* also belong to *s2* in PWN version *v2* by associativity. But Rule 7 produces fallacies when *s1* was split into different target synsets, where *k1* and *k2* are no longer synonyms.

Mapping Rule: *Persistent Synonymy*

$$\& \begin{matrix} Key^{v1}_{k1\in s1} \leftrightarrow Key^{v2}_{k1\in s2} \\ Key^{v1}_{k2\in s1} \end{matrix} \quad \Longrightarrow \quad Key^{v2}_{k2\in s2} \tag{7}$$

Studying the evolution of the sense keys allows us to detect all splits or merges, and to assess their frequency and complexity, i. e. the maximal number of synonym sets involved in one split or merge operation (see Table 4), which permits to precisely identify and count the maximal number of *false positives* that Rule 7 can produce.

2.2.4 Performance analysis

The mappings released by the SKI project (Kafe, 2017b, ski-mappings-pwn) apply only the Mapping Rules 2 and 7. The true performance of these mappings lies somewhere above a *lower bound* that can be calculated by finding the theoretical minimum of the number of correct mapping predictions, and the maximal number of possible fallacies.

	Mapped	**Not mapped**
True	$tp = Keys_{Persist}$	$tn = 0$
False	$fp = Keys_{Removed} \in Synsets_{Persist}$	$fn = Keys_{Removed} \in Synsets_{Removed}$

As reference, we use the imaginary performance of a hypothetic ideal mapping which would be able to map everything accurately, achieving 100% precision and 100% recall. In this ideal situation, there are no true negatives ($tn = 0$), so the sense keys pertaining to the removed synsets from Table 3, which our less ideal mapping cannot map, are false negatives (*fn*). Only the persistent sense keys from Table 2 are the true positives (*tp*), while all the rest of the mapping could be false positives (*fp*). The number of removed keys is equal to the sum of *fp* and *fn*, so *fp* can be obtained either by subtracting *tp* from the length of the mapping, or by subtracting *fn* from the number of $SenseKeys_{removed}$, and both results are expected to be identical, which is verifiable in practice. These values allow us to use standard formulas to calculate *lower bounds* for the *precision* and *recall* of the mappings.

3 Results

3.1 The persistence of word senses

3.1.1 Identity violations in WordNet updates

We started by investigating violation patterns in all WordNet relations and definitions, and found only 239 *attribute violations* in total, when adding all consecutive WordNet updates between the earliest version 1.5 and the latest version 3.1.1, and slightly less (230) in the direct update between the same two versions. The violations peaked during the update from WN 1.6 to WN 1.7, and have declined in absolute numbers since then.

Table 1: Attribute Violations in WordNet Updates

$\mathbf{WN}_{source}$	$\mathbf{WN}_{target}$	*Hypernym*	*Definition*	*Synonym*	*Der./Pertainym*	*Antonym*	*Part holonym*	*Hyponym*	*Member holonym*	*Verb Group*	*Part meronym*	*Derivation*	*Similar to*	TOTAL	ESSENTIAL
1.5	**1.5SC**	15	12	10	4	6	0	1	0	0	0	0	1	*37*	*36*
1.5SC	**1.6**	24	13	8	13	1	0	0	1	0	0	0	0	*43*	*43*
1.6	**1.7**	52	39	16	1	0	8	1	2	2	1	0	0	*66*	*64*
1.7	**1.7.1**	17	10	2	0	0	0	1	0	0	0	0	0	*19*	*18*
1.7.1	**2.0**	26	17	10	3	2	1	2	0	1	0	0	0	*36*	*36*
2.0	**2.1**	8	10	9	1	6	0	0	0	0	0	1	0	*21*	*20*
2.1	**3.0**	8	8	2	0	0	1	0	0	0	0	0	0	*15*	*15*
3.0	**3.1**	0	1	0	0	0	0	0	0	0	0	1	0	*2*	*1*
3.1	**3.1.1**	0	0	0	0	0	0	0	0	0	0	0	0	*0*	*0*
1.5	**3.1.1**	122	45	55	20	10	14	20	3	0	2	0	1	*230*	*214*
1.5	**1.6**	38	19	18	18	7	0	1	1	0	0	0	1	*77*	*76*
1.6	**3.0**	80	72	37	5	7	11	6	2	4	1	0	0	*150*	*144*
3.0	**3.1.1**	0	1	0	0	0	0	0	0	0	0	1	0	*2*	*1*

As a consequence of the low total number of violations, we reviewed all these cases, and found that all the violations of *hypernymy, definitions, synonymy, antonymy, holonymy* and *pertainyms* resulted in genuine violations of semantic identity, which reveals that these relations participate in describing the *essence* of their first argument. Other relations (*hyponymy, verb group, part meronym, derivations, similar to*) were more rarely violated, and only denoted an ESSENTIAL violation when they accompanied a violation of one of the more *essential relations*, but not otherwise. The remaining WordNet relations (like *instance hypernyms*) were never violated, but from the violations of hypernymy which have become instance hypernyms in later WordNet versions (like the *Armstrong* example below), we may infer that *instance hypernymy* is also an *essential relation*, while *instance hyponymy* is not. The full list of *essential violations* found in all consecutive PWN updates is included in Appendix A.

Although the number of identity violations shown in Table 1 was globally low, it is still higher than expected, since essential violations are not allowed in theory. In these results, the relations are ordered from left to right by their decreasing number of violations. In this ordering, hypernymy emerges as the most frequent source of violations, which confirms the importance of hypernymy in expressing essential characteristics of word sense. The TOTALS reflect the cardinality of the set union of all violations, which is lower than their sum, since each violation can manifest itself in several relations simultaneously. Out of these totals, the ESSENTIAL column shows the number of genuine identity violations involving the more *essential relations* mentioned earlier.

The most evident violations of sense identity concern proper names, so the inclusion of named entities in PWN is a major help for recognizing identity breaches, since persons born or dead at different times cannot be the same person, like f. ex. GREGORY%1:18:01, a synonym of pope Gregory I (540?–604) in PWN 2.0 and of Gregory VII (1020–1085) in PWN 2.1. Likewise, places situated in different regions cannot be the same place, like WORCESTER%:1:15:00, which has *Massachusets* as part holonym in PWN 1.6, but *England* in PWN 1.7. The first violation concerns the individual *thisness* of "Gregory", while the second example concerns the shared *whatness* of "Worcester". This shows that WordNet can also express *haecceity* through synonymy, and *quiddity* by subordination to a *holonym*. Likewise, the *quiddity* (astronaut vs. jazzman) of ARMSTRONG%1:18:00 was violated by hypernymy during the same PWN update, while his *haecceity* (Neil vs. Louis) was violated by synonymy.

Violation		ARMSTRONG%1:18:00	ARMSTRONG%1:18:01
Definitions	WN 1.6	**the first man to set foot on the moon (july 20, 1969)**	∄
	WN 1.7	united states jazz trumpeter and bandleader (1900–1971)	united states astronaut; **the first man to set foot on the moon (july 20, 1969)** (1930–)
Hypernyms	WN 1.6	**astronaut%1:18:00** **cosmonaut%1:18:00** **spaceman%1:18:00**	∄
	WN 1.7	cornetist%1:18:00 jazz_musician%1:18:00 jazzman%1:18:00 trumpeter%1:18:00	**astronaut%1:18:00** **cosmonaut%1:18:00** **spaceman%1:18:00**
Synonyms	WN 1.6	**neil_armstrong%1:18:00**	∄
	WN 1.7	louis_armstrong%1:18:00 satchmo%1:18:00	**neil_armstrong%1:18:00**

Common words like SOFT%3:00:05 also present clear examples of such confusions. Here, the WordNet lexicographers changed the sense of SOFT%3:00:05 while assigning the new key SOFT-%3:00:07 to the old sense. The changed part of the sense key is the last field (the *lex_id* of the word *soft*), which was changed from *5* to *7*, inside the lexicographer file *00* (*adj.all*).

Violation		SOFT%3:00:05	SOFT%3:00:07
Definitions	WN 2.0	**used chiefly as a direction or description in music**	∄
	WN 2.1	(of light) transmitted from a broad light source or reflected	**used chiefly as a direction or description in music**
Synonyms	WN 2.0	**piano%3:00:00**	∄
	WN 2.1	diffuse%3:00:00 diffused%3:00:00	**piano%3:00:00**
Antonyms	WN 2.0	**forte%3:00:00** **loud%3:00:02**	∄
	WN 2.1	concentrated%3:00:01 hard%3:00:05	**forte%3:00:00** **loud%3:00:02**

Between WN 1.5 and 3.1.1, the ESSENTIAL violations concerned only 0.1% of the total number of sense keys, so their impact on the global *sense stability* of the PWN keys is almost imperceptible. These low absolute numbers are however only rough estimates and not strict upper bounds, because although this study considers all WordNet relations, it only compared identical definitions. So it is possible that additional violated definitions could be found among the modified definitions that are not accompanied by any relation violation. This eventuality seems less likely, since any *essential violation* in a definition could be expected to entail corresponding changes in the relations. However, 6 out of the 39 violations of identical definitions in the update between PWN 1.6 and 1.7 indeed occurred without any relation violation, which confirms that this eventuality is real. If a similar proportion (6/39 is 15%) holds among all definition violations (including the yet unknown ones), then 85% of these would also present relation violations, and thus already be identified in this study. Then, we may estimate the number of still unknown violations of edited definitions to be approximately close to 15/85 (17.6%) of the number of essential relation violations reported here. Since these numbers are always very low, and only amount to a tiny fraction of the total number of sense keys in any PWN version, the rounded stability percentages presented in the following sections can be expected to hold within very narrow confidence intervals.

3.1.2 Identity preservation in WordNet updates

Table 2 displays the number of persistent, added, and removed sense keys for the nine WordNet updates from version 1.5 to 3.1.1, and four typical long-distance updates between non-consecutive versions, which are relevant for some foreign language wordnets (Dziob, Piasecki, Maziarz, Wieczorek, & Dobrowolska-Pigoń, 2017; Kahusk & Vider, 2017), or studied in previous literature (Daudé et al., 2001; Gonzalez-Agirre et al., 2012; Vossen et al., 2016). Only few wordnets are linked to other versions (Bond & Paik, 2012), so we did not study every possible combination of non-consecutive WordNet versions separately. However, approximate figures can be derived by adding the changes found in the intermediate consecutive versions. For example, a simple addition of the removals between all the versions from WN 2.0 to WN 3.1.1 is sufficient in order to obtain a reliable estimate of this long-distance stability result.

Table 2 shows a high persistence of the sense keys after version 1.6: less than 1% were typically removed between consecutive versions, so the percentage of *persistent* keys was generally above 99. But before version 1.6, the persistence was a little lower, with approx. 3% removals between versions. For long-distance updates, the lost sense keys accumulate: in total 18368 sense keys have been removed since PWN 1.5, so the ratio of keys from PWN 1.5 that persist in the latest PWN

3.1.1 drops to 89.1%. Most often, the number of additions have by far exceeded the deletions, the only exception being the latest WN 3.1.1 update, which mostly consisted in removals.

Overall, the rounded percentages found here are identical with the results reported by Kafe (2017a), which did not consider sense key violations, and the only effect of those is a 0.1% stability decrease in two old PWN updates, which is almost negligible.

3.2 The persistence of synonym sets

Table 3 shows that the synonym sets were always more persistent than the individual sense keys. The lowest persistence rate was 94.4% for the long-distance update from PWN 1.5 to 3.1.1. Again, the overall percentages were identical with Kafe (2017a), except for an almost negligible 0.1% stability decrease in a few PWN updates, after taking the key violations into account.

The superior stability of the synonym sets may actually be expected, considering that removed word senses are mapped to the target synset of their synonyms. For example, although the adjective sense key for "froward" disappeared between WN 3.1 and 3.1.1 because the orthography of the lemma was corrected to "forward", it is still mapped through synonyms like "headstrong". So mappings that link synset offsets have a higher recall than those that only link sense keys, because they cover whole sets of words, and thus avoid some of the losses incurred from the removal of individual sense keys. However, when synsets are split, mapping each key to all its synonyms causes a loss of precision, which we can quantify through a more precise analysis of the splits.

3.2.1 Splits and merges

The following example from PWN 2.1 displays an addition (*medusoid*), a deletion (*medusa#2*), a split (*jellyfish*), and a merge (*medusan*). The deletion of *medusa#2* is implied by the fact that there is already a sense of *medusa* in the target synset.

Sense Key	$\mathbf{WN}_{2.1}$	$\mathbf{WN}_{3.0}$
medusoid%1:05:00::	0	101910252
medusa%1:05:01::	101890584	101910252
medusan%1:05:00::	101891041	101910252
medusa%1:05:02::	101891041	0
jellyfish%1:05:00::	101891041	101910747

The next example shows that the adverb *observably* migrated to its antonym set, during the update from WordNet 2.0 to 2.1. In this case, applying the mapping Rule 7 to its source synonyms *imperceptibly* and *unnoticeably* would aggravate the confusion between synonyms and antonyms, instead of resolving it. To avoid such errors, it is crucial to review all the splits manually. This example also shows that merges do not produce *false positives*, since the other merged source synset (*perceptibly* and *noticeably*) is only mapped to the correct target.

$\textbf{Sense Key}_{split}^{merged}$	$\mathbf{WN}_{2.0}$	$\mathbf{WN}_{2.1}$
imperceptibly%4:02:00::	400369180	400367415
unnoticeably%4:02:00::	400369180	400367415
observably%4:02:00::	400369180	400367669
noticeably%4:02:00::	400369465	400367669
perceptibly%4:02:00::	400369465	400367669

Table 2: Persistence of the SENSE KEYS between WordNet versions

$\mathbf{WN}_{source}$	$\mathbf{WN}_{target}$	$Total_{source}$	$Total_{target}$	*Added*	%	*Removed*	%	*Persist*	%
1.5	**1.5SC**	168082	170243	8502	5	6341	3.8	161741	96.2
1.5SC	**1.6**	170243	173941	9568	5.5	5870	3.4	164373	96.6
1.6	**1.7**	173941	192460	19862	10.3	1343	0.8	172598	99.2
1.7	**1.7.1**	192460	195817	3669	1.9	312	0.2	192148	99.8
1.7.1	**2.0**	195817	203145	9110	4.5	1782	0.9	194035	99.1
2.0	**2.1**	203145	207016	6184	3	2313	1.1	200832	98.9
2.1	**3.0**	207016	206941	2331	1.1	2406	1.2	204610	98.8
3.0	**3.1**	206941	207235	677	0.3	383	0.2	206558	99.8
3.1	**3.1.1**	207235	206353	39	0	921	0.4	206314	99.6
1.5	**3.1.1**	168082	206353	56639	27.4	18368	10.9	149714	89.1
1.5	**1.6**	168082	173941	17894	10.3	12035	7.2	156047	92.8
1.6	**3.0**	173941	206941	40208	19.4	7208	4.1	166733	95.9
3.0	**3.1.1**	206941	206353	716	0.3	1304	0.6	205637	99.4

Table 3: Persistence of the SYNONYM SETS between WordNet versions

$\mathbf{WN}_{source}$	$\mathbf{WN}_{target}$	$Total_{source}$	$Total_{target}$	*Added*	%	*Removed*	%	*Persist*	%
1.5	**1.5SC**	91581	95137	4625	4.9	1341	1.5	90240	98.5
1.5SC	**1.6**	95137	99642	5676	5.7	1236	1.3	93901	98.7
1.6	**1.7**	99642	109377	10004	9.1	406	0.4	99236	99.6
1.7	**1.7.1**	109377	111223	1933	1.7	121	0.1	109256	99.9
1.7.1	**2.0**	111223	115424	4869	4.2	737	0.7	110486	99.3
2.0	**2.1**	115424	117597	3161	2.7	1019	0.9	114405	99.1
2.1	**3.0**	117597	117659	1166	1	1119	1	116478	99
3.0	**3.1**	117659	117791	257	0.2	127	0.1	117532	99.9
3.1	**3.1.1**	117791	117371	15	0	436	0.4	117355	99.6
1.5	**3.1.1**	91581	117371	30350	25.9	5132	5.6	86449	94.4
1.5	**1.6**	91581	99642	10248	10.3	2524	2.8	89057	97.2
1.6	**3.0**	99642	117659	20752	17.6	3025	3	96617	97
3.0	**3.1.1**	117659	117371	273	0.2	563	0.5	117096	99.5

3.2.2 False splits and merges

Many identity violations coincide with a synset being split, merged or both. These interactions reveal that the corresponding splits may only be apparent. For example, from the raw sense keys, it would seem that the synset containing SOFT%3:00:05 was split between WN 2.0 and WN 2.1.

Sense Key	$WN_{2.0}$	ILI	$WN_{2.1}$
diffuse%3:00:00::	0	i6334	301204571
diffused%3:00:00::	0	i6334	301204571
soft%3:00:05::	301408523	i6334	301204571
piano%3:00:00::	301408523	i7979	301510689
soft%3:00:07::	0	i7979	301510689

However, we saw previously that this update violated the integrity of SOFT%3:00:05, which should be mapped to SOFT%3:00:07 in WN 2.1:

Mapping link:

$$WN^{2.0}_{soft\%3:00:05} \leftrightarrow WN^{2.1}_{soft\%3:00:07} \tag{8}$$

Then, if we apply this mapping, the source synset is not split, because the violated key remains a synonym of PIANO. Thus, the number of real splits and merges differs from their apparent number, and depends on the sense key violations. Applying this mapping before other mappings prevents errors that would occur otherwise: if we just map the raw sense keys, the mapping rule 2 discussed earlier will make SOFT%3:00:05 a synonym of DIFFUSE, thus producing a false positive, and break its synonymy with PIANO, resulting in a false negative. Mapping Rule 7 produces a different set of errors (two false positives and zero false negatives), because the synonymy with PIANO would persist, so there would be no false negatives, but an additional false positive would be produced when PIANO also becomes a fallacious synonym of DIFFUSE.

By contrast, the stable synset identifiers from the Inter-Lingual Index (ILI) (Vossen, 2002; Vossen et al., 2016) are not affected by this *false split*, since the ILI-WordNet mapping (UPC TALP, 2017) only maps the involved synsets to their correct targets, without need for any particular handling of the individual word senses. The stability of the ILI-based sense key *soft#i6334* shows that the corresponding PWN key violation could have been avoided by keeping the *lex_id* of SOFT%3:00:05 unchanged as a synonym of PIANO while assigning the new key SOFT%3:00:07 to the synonym of DIFFUSE. So the lexicographers' liberty to freely assign lex_id appears to be the probable cause of the sense key violations found in WordNet.

3.2.3 True splits and merges

In this study, as explained in the Methods section, we blocked the fallacious confusion of violated sense keys, by considering them as removals from the source WN and new additions to the target version. As a consequence, the false splits and merges resulting from sense violations are avoided, and the figures presented in Table 4 correspond to a more restrictive set of true splits and merges. For example, out of 223 *apparent splits* found by Kafe (2017a) for the update from WN 1.6 to 1.7, we only retain 196, which means that the difference, amounting to 27 (8%) *false splits* constitutes an important part (42%) of the 64 ESSENTIAL sense key violations reported in Table 1. So our study shows that, although almost negligible in absolute numbers, the sense key violations have a significant impact on the number of split synsets, which is often approximately 10% lower than reported by Kafe (2017a).

Through all updates, $merged_2$ and $merged_3$ always add up to the total number of merges, so no target synset was ever merged from more than three source synsets. Similarly, after PWN version 1.5SC, $split_2$ and $split_3$ also add up to the total number of splits, so no source synset was split into more than three target synsets. Only in the mapping between WordNet 1.5 and 1.5SC, the total number of splits includes a very small number of four and five-way splits. The number and size of the splits and merges was generally low, and there were always more splits than merges. Almost all splits and merges only involved two synsets, and operations involving three synsets were very rare. Synsets that were split and merged at the same time most often resulted from the migration of a single sense key to another synset.

3.3 The performance of simple sense key mappings

Analysing the sense key-based mappings released by the SKI project (Kafe, 2017b, ski-mappings-pwn) shows, as expected, that applying Mapping Rule 7 increases *recall* but deteriorates *precision* (cf. Table 5). However, after version 1.6, both measures show excellent performance. Compared with Kafe (2017a), the sense key violations have no impact on the *recall* and lead to 0.1% *decreased precision* in only three PWN updates.

This analysis differs from human evaluations by considering the whole PWN dataset, instead of smaller samples, so it provides exact metrics, while human evaluations of limited samples add sample and evaluator biases that can yield higher *standard error*, resulting in wider confidence intervals. Larger human evaluations are needed, as well as deeper analyses, since both approaches have complementary merits, and allow meaningful comparisons.

Only few partial studies have previously been conducted about the performance of mappings between PWN versions. Daudé et al. (2001) produced a complete *synset offset* mapping from PWN 1.5 to 1.6, by applying a relaxation labelling algorithm, with a set of constraints that involved all semantic relations, and additional heuristics such as gloss similarity. They evaluated the results manually, by applying different constraint sets on samples drawn from the monosemous vs. ambiguous nouns, verbs, adjectives and adverbs (4200 synsets in total), and found 98.8% precision and 98.9% recall for the nouns overall, when using the complete constraint set. In all cases, recall was higher than precision. A particular strength of these mappings is their ability to correct identity violations, which constitute a weakness for sense keys.

By comparison, the corresponding sense key-based mapping for this old PWN update also shows higher *lower bounds* for recall (97.6%) than for precision (95%). However, with the later versions this tendency was inversed, since precision was consistently higher than recall, and both figures stayed mostly above 99.5% in the later consecutive PWN mappings.

4 Discussion

4.1 WordNet synsets are very stable

After establishing that the number of *sense key violations* in WordNet was almost negligible, we simply followed the stable sense keys between all WordNet versions, and saw that the synonym sets have remained very stable throughout every update. There was never more than a few hundred split or merged synonym sets between consecutive versions and, after version 1.6, the complexity of these changes was often the lowest possible, because each split or merge almost always involved only two synsets, and never more than three.

This is the first comprehensive analysis of the persistence of semantic identity in all the WordNet versions released between 1995 and 2012. Our results show that both existing mapping paradigms (the recommended sense-key based, and the dominant synset-based) have their respective specific challenge (the identity violations vs. the split synsets), which is often a strength of the other paradigm. Thus, all WordNet mappings can be improved by focusing on their specific sources of errors. In both cases, the number of problems is limited, which indicates that it will become possible to update wordnets to new PWN versions with greater confidence and less effort. So, alt-

Table 4: Splits and Merges in the SYNONYM SETS between WordNet versions

$\mathbf{WN}_{source}$	$\mathbf{WN}_{target}$	*Split*	$split_2$	$split_3$	*Merged*	$merged_2$	$merged_3$	$\&^{merged}_{split}$
1.5	**1.5SC**	472	443	25	226	217	9	134
1.5SC	**1.6**	252	238	14	199	197	2	90
1.6	**1.7**	196	192	4	63	63	0	37
1.7	**1.7.1**	53	52	1	20	20	0	6
1.7.1	**2.0**	116	113	3	50	50	0	24
2.0	**2.1**	81	77	4	54	54	0	17
2.1	**3.0**	79	78	1	64	63	1	26
3.0	**3.1**	33	33	0	31	31	0	11
3.1	**3.1.1**	1	1	0	0	0	0	0
1.5	**3.1.1**	1120	1048	67	613	599	14	326
1.5	**1.6**	701	652	44	408	396	12	223
1.6	**3.0**	506	489	17	231	229	2	107
3.0	**3.1.1**	33	33	0	31	31	0	11

Table 5: Performance *lower bounds* of sense-key mappings between WordNet versions

$\mathbf{WN}_{source}$	$\mathbf{WN}_{target}$	*tp*	*fp*	*fn*	*Precision*	*Recall*
1.5	**1.5SC**	161741	4141	2200	97.5	98.7
1.5SC	**1.6**	164373	4173	1697	97.5	99
1.6	**1.7**	172598	796	547	99.5	99.7
1.7	**1.7.1**	192148	156	156	99.9	99.9
1.7.1	**2.0**	194035	735	1047	99.6	99.5
2.0	**2.1**	200832	711	1602	99.6	99.2
2.1	**3.0**	204610	730	1676	99.6	99.2
3.0	**3.1**	206558	180	203	99.9	99.9
3.1	**3.1.1**	206314	52	869	100	99.6
1.5	**3.1.1**	149714	10256	8112	93.6	94.9
1.5	**1.6**	156047	8170	3865	95	97.6
1.6	**3.0**	166733	2632	4576	98.4	97.3
3.0	**3.1.1**	205637	232	1072	99.9	99.5

hough rarely studied, the persistence of semantic identity is an essential question that has acute consequences for the interoperability of all projects that use WordNet data.

Lexicographers can use Tables 1, 2, 3, 4 and 5 to estimate the effort required to update a resource between two PWN versions. For example, when updating to PWN 3.0, a resource that uses PWN 1.6 sense keys and just applies Rule 2 would obtain almost perfect precision (subtracting the 144 ESSENTIAL violations from Table 1), and 95.9% recall (Table 2), which can be improved by a review of the 7208 removed sense keys, as well as the eventual collapses of identical words (like *medusa#2* in a previous example) resulting from the 231 merged synsets (Table 4). Mapping Rule 7 improves recall (97.3% in Table 5), which can be further improved by reviewing the same 231 merges, and the 4576 *false negatives* remaining from the 7208 removed sense keys that belong to the 3025 *removed* synsets (Table 3), while the rest of these 7208 removed sense keys could be *false positives* produced by Rule 7, and need to also be reviewed in order to increase *precision*, in addition to the 506 splits, which cause a decrease of *precision* that does not affect sense keys.

So these results confirm that "sense keys are the best way to represent a sense" (WordNet-team, 2010, Senseidx), but only by a small margin. Contrary to expectations, synset identifiers provide a reasonable alternative, since the splits between most versions are relatively few and simple. As a consequence, stable synset identifiers like the Inter-Lingual Index (ILI) appear viable, although they will need to be complemented by mapping links between ILI-based sense keys, in order to handle the split synsets.

4.2 Practical application

For the older wordnets that are still mapped to PWN 1.5, like Polish (Dziob et al., 2017) and Estonian (Kahusk & Vider, 2017), upgrading to PWN 3.1.1 requires to review the intersection of the source data with the 1120 PWN splits reported in Table 4, and the 214 sense key violations counted in Table 1, and listed in Appendix A. More recent projects linked to PWN 2.0 are in a luckier position, since the addition of all the changes that have occurred in the intermediate WN versions yields only 194 splits, and 36 violations.

Obviously, projects already linked to WN 3.0 enjoy a more fortunate situation. For example, updating the wordnets from MCR30-2016 (Gonzalez-Agirre et al., 2012) to PWN 3.1 is much simpler, since only 33 splits need to be checked. One of these is the following example from WordNet, where *Pluto* was moved from the Greek to the Roman "gods of the underworld", while the corresponding Greek name *Plouton* is not in PWN yet.

Sense Key	$\mathbf{WN}_{3.0}$	$\mathbf{ILI}_{3.0}$	$\mathbf{ILI}_{3.1}$	$\mathbf{WN}_{3.1}$
aides%1:18:00::	109570298	i86957	i86957	109593427
aidoneus%1:18:00::	109570298	i86957	i86957	109593427
hades%1:18:00::	109570298	i86957	i86957	109593427
pluto%1:18:00::	109570298	i86957	i86958	109593643
dis%1:18:00::	109570522	i86958	i86958	109593643
orcus%1:18:00::	109570522	i86958	i86958	109593643
dis_pater%1:18:00::	0	0	i86958	109593643

The ILI 3.1 mapping (GWA, 2017) provides correct identifiers at the synset level, but cannot help in mapping local translations of *Pluto* to their adequate PWN 3.1 synset, so the eventual local splits have to be resolved by local lexicographers. For example, the Spanish WordNet from MCR30-2016 (Gonzalez-Agirre et al., 2012) also includes the involved synsets. Thus, the Spanish lexicographers need to consider whether *Plutón* corresponds to either the Greek or the Latin name, or eventually to both, and other WN 3.0-based resources also face the same issue. In this particular example, applying Mapping Rule 7 would place *Plutón* in both the Roman and the Greek synsets, which seems adequate if both gods are the same.

Persistent synset identifiers like ILI are useful at the coarse synonym-set level, but they need to be complemented by sense keys to handle the more precise word-level. Since words are unique within each synset, a sense key can be constructed by simply combining each word and its synset identifier. If persistent ILI identifiers are used for the synsets, the corresponding ILI-based sense keys are also persistent, except when they denote one of the few violated or migrated senses. Then, a mapping link between ILI-based sense keys can handle these exceptional cases, and for example express the migration of *Pluto* to a different synset. This format can express mappings between any source and target word senses, and is thus also able to adequately to handle the violations of semantic identity listed in Appendix A.

Mapping ILI-based sense keys:

$$WN^{3.0}_{pluto\#i86957} \leftrightarrow WN^{3.1}_{pluto\#i86958} \tag{9}$$

4.3 Strengths and limitations

Objectively, the key violation mappings from Appendix A provide only 0.1% increased *precision* in just a few PWN versions. But this tiny quantitative gain may correspond to a larger *qualitative* improvement of the concerned PWN links, since the more severe violations (like confusing clearly distinct cities or persons) have a damaging impact on the perceived quality of a lexical resource, and the confidence in its use.

This study relies crucially on the low number of identity violations in WordNet. But, as mentioned earlier, we do not know whether all sense key violations were identified here, so the results in Table 1 are only estimates. Although the PWN Sensemap (WordNet-team, 2010, Sensemap) could be expected to map additional violations, this does not seem to be the case: on the contrary, applying our general formula for finding sense key violations (5) in Sensemap just reveals a few dubious mappings that could be avoided by considering the definitions. For the moment, we may estimate that the real violations are probably too few to change the rounded percentages reported here, but this will need to be confirmed in future studies. Some controversy must be expected, since we still miss a consensual and exhaustive definition of *semantic essence*, so diverging appreciations of the notion of *semantic identity* are inevitable.

Also, the present study is limited to only two primary *mapping inference rules*, based on *sense key identity* (2) and *persistent synonymy* (7). Additional mapping links can also be inferred automatically from gloss similarity and other relations, as in (Daudé et al., 2001). However, since these additional heuristics are more uncertain, they should be studied separately, and applied at a later stage.

Our results show that synset identifiers like the ILI could be stable in theory, but the actual stability of the ILI has not yet been investigated in practice. A similar study of the various ILI versions would be interesting.

Last but not least, we still do not know how to avoid future identity violations, both in PWN and in other wordnets. According to Christiane Fellbaum[2], "The WordNet lexicographers are free to change the sense inventory as they see fit", though in later versions, the PWN compiler (WordNet-team, 2010, grind) flags eventual duplicate *lex_ids* for the same word within a lexical file. According to Randee Tengi[2], the *grind* program then "leaves it to the lexicographer to view the synsets and be sure that the correct *lex_id* is used to carry forward a synset with the same meaning that it had in the previous version". This ensures *unique* sense keys, but does not protect against identity violations. Perhaps a closer study of the examples in Appendix A can lead to a better understanding of this potential pitfall.

[2]Personal communication.

5 Conclusion

We followed the *sense keys* between WordNet versions, and obtained exact figures for the number of *added* and *removed* word senses and synonym sets, as well as the number and complexity of the *split* and *merged* synsets. We found that the splits and merges between versions were few and simple, and that the synsets have remained very stable throughout. Even though their identifiers are unstable, the synsets were always more persistent than the sense keys, especially in the earlier versions. However the sense keys have the advantage of almost perfect precision, and have stayed almost as persistent as the synsets after PWN 1.6. So both identifiers provide almost equivalent support for highly accurate mappings between the later WordNet versions.

We can thus answer an interrogation from the *CfWNs 2017* workshop discussions (Bond, 2017): yes, *wordsense keys* are needed in order to ensure integrity and precision in the lexical database, but these keys could have other formats than the old PWN sense keys. For example, new sense keys based on ILI identifiers show particular promises because of their superior handling of the *false splits* induced by violations of some PWN sense keys. Concerning another discussion topic, our first result showed that violations of identity were easiest to recognize with proper names, which advocates for including named entities in wordnets, although eventually in a separate database.

When relying on the solid baseline provided by the persistent sense keys and synsets, the lexicographic work required to update synset-mapped resources to newer versions of WordNet can essentially be reduced to a manual review of the exceptional identity violations, relatively few splits and merges, and a moderate amount of removals. The burden would be further reduced if identity violations could be avoided in future WordNet versions.

This study was only possible because PWN offers permanent sense keys, so we may expect that other wordnets with permanent identifiers also enjoy more accurate traceability, leading to enhanced interoperability.

References

Bond, F. (2017). Discussion notes. In J. P. McCrae, F. Bond, P. Buitelaar, P. Cimiano, T. Declerck, J. Gracia, I. Kernerman, E. Montiel Ponsoda, N. Ordan, & M. Piasecki (Eds.), *Proceedings of the LDK 2017 Workshops: 1st Workshop on the OntoLex Model (OntoLex-2017), Shared Task on Translation Inference Across Dictionaries & Challenges for Wordnets co-located with 1st Conference on Language, Data and Knowledge (LDK 2017)*. Retrieved August 31, 2017, from `http://ceur-ws.org/Vol-1899/`

Bond, F., & Paik, K. (2012). A survey of wordnets and their licenses. In P. Vossen & C. Fellbaum (Eds.), *Proceedings of the 6th Global WordNet Conference, Matsue, Japan*. Brno: Tribun EU.

Bond, F., Fellbaum, C., Hsieh, S.-K., Huang, C.-R., Pease, A., & Vossen, P. (2014). A multilingual lexico-semantic database and ontology. In P. Buitelaar & P. Cimiano (Eds.), *Towards the Multilingual Semantic Web* (pp. 243–258). Springer.

Chen, P. (1976). The entity relationship model — toward a unified view of data. *ACM Transactions on Database Systems*, *1*(1), 312–339. `https://doi.org/10.1145/320434.320440`

Cohen, S. M. (2016). Aristotle's metaphysics. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University. Retrieved from `https://plato.stanford.edu/entries/aristotle-metaphysics/`

Cross, R. (2014). Medieval theories of haecceity. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University. Retrieved from `https://plato.stanford.edu/entries/medieval-haecceity/`

Daudé, J., Padró, L., & Rigau, G. (2001). A complete wn1.5 to wn1.6 mapping. In *Proceedings of the NAACL Workshop "WordNet and Other Lexical Resources: Applications, Extensions and Customizations" (NAACL'2001)*. Pittsburg, PA.

Dziob, A., Piasecki, M., Maziarz, M., Wieczorek, J., & Dobrowolska-Pigoń, M. (2017). Towards revised system of verb wordnet relations for Polish. In J. P. McCrae, F. Bond, P. Buitelaar, P. Cimiano, T. Declerck, J. Gracia, I. Kernerman, E. Montiel Ponsoda, N. Ordan, & M. Piasecki (Eds.), *Proceedings of the LDK 2017 Workshops: 1st Workshop on the OntoLex Model (OntoLex-2017), Shared Task on*

Translation Inference Across Dictionaries & Challenges for Wordnets co-located with 1st Conference on Language, Data and Knowledge (LDK 2017) (pp. 174–187). Retrieved August 31, 2017, from `http://ceur-ws.org/Vol-1899/`

Fellbaum, C. (Ed.). (1998). *WordNet: An electronic lexical database.* Cambridge, MA: MIT Press.

Gonzalez-Agirre, A., Laparra, E., & Rigau, G. (2012). Multilingual central repository version 3.0: Upgrading a very large lexical knowledge base. In P. Vossen & C. Fellbaum (Eds.), *Proceedings of the 6th Global WordNet Conference, Matsue, Japan.* Brno: Tribun EU.

GWA. (2017). ili-map-pwn31.tab. In *Collaborative Inter-Lingual Index (CILI).* GitHub. Retrieved April 15, 2017, from `https://www.github.com/globalwordnet/ili`

Kafe, E. (2012). Wordnet mapping. In *HyperDic hyper-dictionary.* MegaDoc. Retrieved April 15, 2017, from `http://www.hyperdic.net/en/doc/mapping`

Kafe, E. (2017a). How stable are WordNet synsets? In J. P. McCrae, F. Bond, P. Buitelaar, P. Cimiano, T. Declerck, J. Gracia, I. Kernerman, E. Montiel Ponsoda, N. Ordan, & M. Piasecki (Eds.), *Proceedings of the LDK 2017 Workshops: 1st Workshop on the OntoLex Model (OntoLex-2017), Shared Task on Translation Inference Across Dictionaries & Challenges for Wordnets co-located with 1st Conference on Language, Data and Knowledge (LDK 2017)* (pp. 113–124). Retrieved August 31, 2017, from `http://ceur-ws.org/Vol-1899/`

Kafe, E. (2017b). Sense key index for inter-operability between wordnet-related projects. In *SKI.* GitHub, Retrieved April 25, 2017, from `https://www.github.com/ekaf/ski`

Kahusk, N., & Vider, K. (2017). The revision history of Estonian Wordnet. In J. P. McCrae, F. Bond, P. Buitelaar, P. Cimiano, T. Declerck, J. Gracia, I. Kernerman, E. Montiel Ponsoda, N. Ordan, & M. Piasecki (Eds.), *Proceedings of the LDK 2017 Workshops: 1st Workshop on the OntoLex Model (OntoLex-2017), Shared Task on Translation Inference Across Dictionaries & Challenges for Wordnets co-located with 1st Conference on Language, Data and Knowledge (LDK 2017)* (pp. 164-173). Retrieved August 31, 2017, from `http://ceur-ws.org/Vol-1899/`

Navigli, R. & Ponzetto, S. P. (2010). Babelnet: Building a very large multilingual semantic network. In *ACL'10: Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, Uppsala, Sweden, 11–16 July 2010* (pp. 216–225). Stroudsburg, PA: Association for Computational Linguistics.

Niles, I., & Pease, A. (2003). Linking lexicons and ontologies: Mapping wordnet to the suggested upper merged ontology. In *Proceedings of The 2003 International Conference on Information and Knowledge Engineering (IKE 03), Las Vegas.* (pp. 412–416). Association for Computational Linguistics.

Piasecki, M., Broda, B., & Szpakowicz, S. (2009). *A wordnet from the ground up.* Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej.

R-team. (2017). R version 3.3.3. In *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. Retrieved January 12, 2018, from `https://www.R-project.org/`

Suchanek, F. M., Kasneci, G., & Weikum, G. (2008). Yago: A large ontology from Wikipedia and WordNet. *Web Semantics: Science, Services and Agents on the World Wide Web*, *6*(3), 203–217. `https://doi.org/10.1016/j.websem.2008.06.001`

UPC TALP. (2017). Older wordnet mappings. In *Collaborative Inter-Lingual Index (CILI).* Global Wordnet Association. Retrieved January 12, 2018, from `https://www.github.com/globalwordnet/ili`

Vossen, P. (2002). *EuroWordNet general document.* EWN.

Vossen, P., Bond, F., & McCrae, J. P. (2016). Toward a truly multilingual Global Wordnet Grid. In V. Barbu Mititelu, C. Forascu, C. Fellbaum, & P. Vossen (Eds.), *Proceedings of the Eighth Global WordNet Conference* (pp. 419–426). Retrieved from `http://jiangbian.me/papers/2016/gwc2016.pdf`

WordNet-team. (2010). Wordnet 3.0 reference manual. In *WordNet Documentation.* Princeton University. Retrieved from `https://wordnet.princeton.edu/documentation`

A Mapping of the Identity Violations in WordNet

WN 1.5 to WN 1.5SC: *alveolar* %3:01:00 → 3:01:02, *anime* %1:27:00 → 1:27:01, *antiphonal* %3:01:00 → 3:01:01, *attentiveness* %1:07:01 → 1:07:00, *butt* %1:06:01 → 1:06:04, *ceylon* %1:15:-

00 → 1:15:01, *cheese* %1:13:00 → 1:13:01, *closely* %4:02:00 → 4:02:01, *conterminous* %5:00:00:-connected:00 → 5:00:01:connected:00, *countrywoman* %1:18:00 → 1:18:01, *cul_de_sac* %1:06:00 → 1:06:01, *decisively* %4:02:01 → 4:02:00, *evidence* %1:10:01 → 1:10:00, *extended* %3:00:00 → 3:00:04, *fugitive* %1:18:00 → 1:18:01, *gambler* %1:18:00 → 1:18:01, *hieratic* %3:01:00 → 3:01:01, *indecisively* %4:02:00 → 4:02:01, *modality* %1:24:00 → 1:24:01, *normative* %5:00:00:standard:02 → 5:00:00:standard:03, *porto_rico* %1:15:00 → 1:15:01, *prodigally* %4:02:00 → 4:02:01, *psychologically* %4:02:00 → 4:02:01, *puerto_rico* %1:15:00 → 1:15:01, *purple* %5:00:00:colored:00 → 5:00:00:-chromatic:00, *raise* %2:32:00 → 2:32:01, *running* %3:00:00 → 3:00:01, *secluded* %5:00:00:private:00 → 5:00:02:private:00, *sliver* %1:17:00 → 1:17:01, *tatar* %1:18:00 → 1:18:01, *topsy-turvy* %4:02:00 → 4:02:01, *undisturbed* %3:00:02 → 3:00:04, *unoffending* %3:00:00 → 3:00:02, *urge_on* %2:32:00 → 2:32:01, *yard* %1:15:00 → 1:06:02, *yell* %2:32:00 → 2:32:01

WN 1.5SC to WN 1.6: *alyssum* %1:20:00 → 1:20:01, *benedict* %1:18:00 → 1:18:01, *book* %2:41:01 → 2:41:03, *caesarian* %3:01:00 → 3:01:01, *canalisation* %1:04:00 → 1:04:01, *canalization* %1:04:00 → 1:04:01, *capsid* %1:05:00 → 1:05:01, *casuistical* %3:01:00 → 3:01:01, *casuistic* %3:01:01 → 3:01:00, *chloride* %1:27:00 → 1:27:01, *condense* %2:30:05 → 2:30:00, *cornhusker* %1:18:00 → 1:18:01, *crack* %2:30:09 → 2:30:01, *cut* %2:38:01 → 2:38:15, *cystic* %3:01:00 → 3:01:01, *deeply* %4-:02:03 → 4:02:00, *deliriously* %4:02:00 → 4:02:01, *deliriously* %4:02:01 → 4:02:00, *extend* %2:30:01 → 2:30:09, *gin* %1:06:00 → 1:06:01, *hellenic* %3:01:00 → 3:01:01, *house* %1:06:00 → 1:06:01, *house* %1:06:00 → 1:06:02, *invasive* %3:00:00 → 3:00:01, *latitude* %1:15:00 → 1:15:01, *liberally* %4:02:00 → 4:02:01, *melodramatically* %4:02:00 → 4:02:01, *melodramatically* %4:02:01 → 4:02:00, *nationally* %4:02:00 → 4:02:02, *nationally* %4:02:02 → 4:02:00, *olive* %1:20:01 → 1:20:02, *pleomorphism* %-1:19:00 → 1:19:01, *repeater* %1:18:00 → 1:18:01, *residuary* %3:01:00 → 3:01:01, *school* %1:14:00 → 1:14:03, *scup* %1:05:02 → 1:05:01, *smoke* %1:06:00 → 1:06:01, *surgery* %1:06:00 → 1:06:01, *teardrop* %1:08:00 → 1:25:00, *unaccented* %5:00:00:unstressed:00 → 3:00:04, *visual_image* %1:09-:00 → 1:09:01, *yam* %1:13:00 → 1:13:02, *yam* %1:13:02 → 1:13:00

WN 1.6 to WN 1.7: *apostle* %1:18:01 → 1:18:02, *armstrong* %1:18:00 → 1:18:01, *bolshevik* %1:18:00 → 1:18:01, *buster* %1:18:00 → 1:18:04, *coiner* %1:18:00 → 1:18:02, *cooper* %1:18:00 → 1:18:03, *corchorus* %1:20:00 → 1:20:01, *corydalis* %1:20:00 → 1:20:01, *c* %1:10:00 → 1:10:01, *don-_juan* %1:18:00 → 1:18:01, *dylan* %1:18:00 → 1:18:01, *egress* %1:11:00 → 1:11:01, *flit* %1:04:00 → 1:04:01, *gilbert* %1:18:00 → 1:18:02, *gilbert* %1:18:00 → 1:18:03, *goliath* %1:18:00 → 1:18:01, *gregory* %1:18:02 → 1:18:03, *hair* %1:08:00 → 1:08:01, *hanoverian* %1:18:00 → 1:18:01, *hertz* %1:18:00 → 1:18:01, *homeboy* %1:18:00 → 1:18:01, *humor* %1:08:00 → 1:08:01, *humour* %1:08:00 → 1:08:01, *interception* %1:04:00 → 1:04:01, *jazz* %1:10:00 → 1:10:02, *judicially* %4:02:00 → 4:02:01, *lag* %-2:35:00 → 2:35:01, *lister* %1:18:00 → 1:18:01, *lumper* %1:18:00 → 1:18:01, *macedonia* %1:15:00 → 1:15:01, *madagascar* %1:15:00 → 1:15:01, *manikin* %1:18:00 → 1:18:01, *mannikin* %1:18:00 → 1:18:01, *marshall_islands* %1:15:00 → 1:15:01, *marshall_islands* %1:15:01 → 1:15:00, *micronesia* %1:15:00 → 1:15:01, *normality* %1:26:00 → 1:07:01, *notch* %1:06:00 → 1:25:00, *pestle* %1:06:00 → 1:06:02, *pestle* %1:06:01 → 1:06:00, *peter_pan* %1:18:00 → 1:18:01, *philistine* %1:18:00 → 1:18:01, *pilot* %1:06:00 → 1:06:01, *play* %2:36:11 → 2:36:13, *plunger* %1:18:01 → 1:18:00, *recollection* %-1:09:01 → 1:09:02, *romanticise* %2:30:00 → 2:31:00, *saul* %1:18:00 → 1:18:01, *scribbler* %1:18:00 → 1:18:01, *seth* %1:18:00 → 1:18:01, *simpson* %1:18:00 → 1:18:01, *solomon_islands* %1:15:00 → 1:15:01, *solomon_islands* %1:15:01 → 1:15:00, *spark_gap* %1:06:00 → 1:06:01, *subversion* %1:-04:00 → 1:04:01, *subvert* %2:41:00 → 2:41:01, *supposal* %1:09:00 → 1:09:01, *tabernacle* %1:06:00 → 1:06:01, *tramcar* %1:06:00 → 1:06:01, *tuvalu* %1:15:00 → 1:15:01, *tuvalu* %1:15:01 → 1:15:00, *victimise* %2:41:00 → 2:41:01, *william_gilbert* %1:18:00 → 1:18:01, *worcester* %1:15:00 → 1:15:01

WN 1.7 to WN 1.7.1: *anastomose* %2:35:00 → 2:35:01, *annexation* %1:04:00 → 1:04:01, *ball* %1:14:00 → 1:11:00, *constable* %1:18:01 → 1:18:02, *county* %1:15:00 → 1:15:01, *davis* %1:18:04 → 1:18:05, *detribalisation* %1:04:00 → 1:04:01, *detribalization* %1:04:00 → 1:04:01, *respiration* %1:04-:00 → 1:04:01, *respiration* %1:04:00 → 1:04:02, *revelation* %1:10:02 → 1:10:01, *reversion* %1:11:00

→ 1:11:01, *spider's_web* %1:06:00 → 1:06:01, *spider_web* %1:06:00 → 1:06:01, *stuff* %2:30:02 → 2:30:12, *transposition* %1:11:00 → 1:11:01, *vaticinate* %2:32:00 → 2:32:01, *vegetate* %2:29:00 → 2:29:01

WN 1.7.1 to WN 2.0: *ali* %1:18:00 → 1:18:01, *antagonistic* %3:00:00 → 3:00:02, *antagonistic* %3:00:02 → 3:00:01, *book_of_psalms* %1:10:00 → 1:10:01, *carboniferous* %3:01:00 → 3:01:02, *carriage* %1:06:00 → 1:06:03, *chartist* %1:18:00 → 1:18:01, *check_out* %2:40:00 → 2:40:02, *co* %1:27:00 → 1:27:01, *du_maurier* %1:18:00 → 1:18:01, *eden* %1:15:00 → 1:09:00, *entropy* %1:-07:00 → 1:07:01, *featherweight* %1:18:00 → 1:18:02, *heavyweight* %1:18:01 → 1:18:04, *hold_over* %2:42:00 → 2:42:02, *indorse* %2:41:00 → 2:41:02, *jehad* %1:04:00 → 1:04:01, *jihad* %1:04:00 → 1:04:01, *lachrymal* %3:01:00 → 3:01:01, *lacrimal* %3:01:00 → 3:01:01, *lessing* %1:18:00 → 1:18:01, *light_heavyweight* %1:18:00 → 1:18:02, *lightweight* %1:18:00 → 1:18:03, *mew* %2:32:00 → 2:32:03, *middleweight* %1:18:00 → 1:18:02, *network* %1:06:00 → 1:06:02, *network* %1:06:00 → 1:06:03, *net* %1:06:01 → 1:06:00, *pragmatist* %1:18:00 → 1:18:01, *quackery* %1:04:00 → 1:04:01, *quantum* %1:-09:00 → 1:09:01, *reciprocal* %1:24:00 → 1:24:01, *reseau* %1:06:00 → 1:06:01, *ship_builder* %1:18:00 → 1:18:01, *velocipede* %1:06:01 → 1:06:02, *welterweight* %1:18:00 → 1:18:02

WN 2.0 to WN 2.1: *ab* %1:08:00 → 1:08:01, *calisthenics* %1:04:00 → 1:04:01, *callisthenics* %1:04:00 → 1:04:01, *dead* %3:00:02 → 3:00:01, *dry* %3:00:05 → 3:00:06, *eulogy* %1:10:00 → 1:10-:01, *forwardness* %1:07:00 → 1:07:02, *gregory* %1:18:01 → 1:18:00, *gregory* %1:18:02 → 1:18:01, *iconoclast* %1:18:00 → 1:18:01, *live* %3:00:00 → 3:00:01, *margin* %1:07:00 → 1:07:01, *militant* %5:00:00:unpeaceful:00 → 5:00:01:unpeaceful:00, *murkily* %4:02:00 → 4:02:01, *put_up* %2:40:00 → 2:40:01, *rough* %3:00:04 → 3:00:05, *single* %3:00:00 → 3:00:05, *soft* %3:00:05 → 3:00:07, *sport* %1:18:02 → 1:18:03, *true_lover's_knot* %1:06:00 → 1:06:01

WN 2.1 to WN 3.0: *assyrian* %1:10:00 → 1:10:01, *clean* %5:00:00:perfect:00 → 5:00:00:-easy:01, *dance_music* %1:10:00 → 1:10:01, *floor* %1:17:00 → 1:17:02, *gratefully* %4:02:00 → 4:-02:01, *gravy* %1:13:00 → 1:13:01, *incessantly* %4:02:00 → 4:02:02, *overnight* %4:02:00 → 4:02:01, *predetermination* %1:09:00 → 1:09:01, *reciprocative* %5:00:00:reciprocal:00 → 5:00:01:reciprocal-:00, *substance* %1:03:00 → 1:27:00, *superbug* %1:05:00 → 1:05:01, *titular* %3:01:02 → 3:01:00, *verbalisation* %1:10:00 → 1:10:01, *verbalization* %1:10:00 → 1:10:01

WN 3.0 to WN 3.1: *decadent* %5:00:00:indulgent:00 → 5:00:00:immoral:00

Acknowledgment
This article expands a paper Kafe (2017a) presented at the *Challenges for Wordnets* workshop in Galway, Ireland (CfWns17), and benefited from substantial constructive suggestions by the anonymous reviewers. Special thanks to Christiane Fellbaum and Randee Tengi for helpful information and comments about WordNet.
The work was funded at the author's own expense.

The author declares that he has no competing interests.

Publisher: Institute of Slavic Studies, Polish Academy of Sciences, University of Silesia & The Slavic Foundation