Acquiring a Poor Man's Inflectional Lexicon for German

Peter Adolphs

Deutsches Forschungszentrum für Künstliche Intelligenz, Projektbüro Berlin
Rosenstraße 2
D-10178 Berlin
peter.adolphs@dfki.de

Abstract

Many NLP modules and applications require the availability of a module for wide-coverage inflectional analysis. One way to obtain such analyses is to use an morphological analyser in combination with an inflectional lexicon. Since large text corpora nowadays are easily available and inflectional systems are in general well understood, it seems feasible to acquire lexical data from raw texts, guided by our knowledge of inflection. I present an acquisition method along these lines for German. The general idea can be roughly summarised as follows: first, generate a set of lexical entry hypotheses for each word-form in the corpus; then, select hypotheses that explain the word-forms found in the corpus "best". To this end, I have turned an existing morphological grammar, cast in finite-state technology (Schmid et al., 2004), into a hypothesiser for lexical entries. Irregular forms are simply listed so that they do not interfere with the regular rules used in the hypothesiser. Running the hypothesiser on a text corpus yields a large number of lexical entry hypotheses. These are then ranked according to their validity with the help of a statistical model that is based on the number of attested and predicted word forms for each hypothesis.

1. Introduction

Many NLP modules and applications require the availability of a module for wide-coverage inflectional analysis. Such a module should enable the user to determine the set of all possible inflectional analyses for a given word form. One way to provide these analyses is to look up the word form in a full-form lexicon. A better solution in terms of maintainability is to use an inflectional lexicon, which lists inflectional stems with their inflectional classes, combined with a morphological analyser that uses those lexical entries to analyse the word forms. To my knowledge, there is no such module freely available for German[1]. Furthermore, existing inflectional lexicons constantly need to be expanded, for instance, to encompass domain-specific vocabulary.

The manual creation and maintenance of an inflectional lexicon is a dull and strenuous task. Since large text corpora nowadays are easily available and inflectional systems are in general well understood, it seems feasible to acquire lexical data from raw texts, guided by our knowledge of inflection. Several methods along these lines have been developed in recent years for different languages including Russian (Oliver et al., 2003), Croatian (Oliver & Tadić, 2004), French (Clément et al., 2004; Forsberg et al., 2006), and Slovak (Sagot, 2005).

I present an acquisition method for German that takes only the tokens of a corpus and their frequencies as input, and produces a list of ranked lexical entry candidates as output. Lexical entries comprise the inflectional stem and the inflectional class of a lexical word (lexeme). They serve as a basis for inflectional analysis, where word forms are mapped to the set of possible combinations of lexical identifiers and morphosyntactic properties, as well as generation, where lexical representations are mapped back to word forms. The method can be summarised as follows: first, generate a set of lexical entry hypotheses for each word-form in the corpus; then, select hypotheses that explain the word-forms found in the corpus "best". To this end, I have turned an existing morphological grammar, cast in finite-state technology, into a hypothesiser for lexical entries. Irregular forms are simply listed, so they do not interfere with the regular rules used in the hypothesiser. Running the hypothesiser on a text corpus yields a large number of lexical entry hypotheses. These are then ranked according to their validity with the help of a statistical model that is based on the number of attested and predicted word forms for each hypothesis.

[1] *Morphy* (Lezius, 2000) is a morphological analyser for German covering inflection and a limited set of compounding, and can be used at no charge. Unfortunately, it only provides a graphical user interface which cannot be embedded into non-interactive software pipelines. The tool offers a function to export its lexical data as a full-form lexicon instead.

2. Acquisition System

2.1. System Architecture

The acquisition system described in the following operates on a tokenised text corpus, which is treated as a 'bag of words', that is, the context of particular tokens is not considered in this work. It is assumed that the tokeniser splits the text at white space and at punctuation marks. Further preprocessing steps such as a normalisation of sentence-initial upper-case letters or an undoing of hyphenation have not been applied. However, in order to keep the set of problematic forms small, I have excluded tokens with non-alphabetic characters from the acquisition process.

The task of the system is to propose new lexical entries for open-class words. All word forms are therefore first analysed by a lexicon-based inflectional analyser, and only unknown word forms are considered for finding candidates for new lexical entries. The actual acquisition process is then realised as a two-step procedure:

1. generation of lexical entry hypotheses according to a computational model of German inflection for all unknown word forms of the corpus

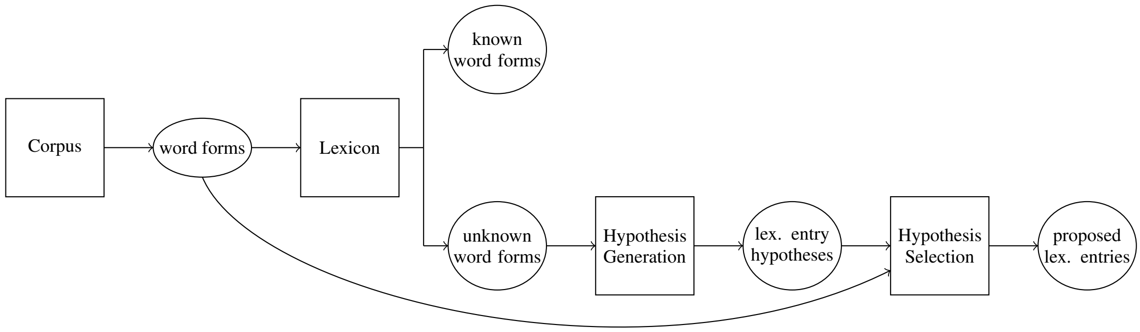

Figure 1: Overview of the acquisition system

2. selection of plausible lexical entry hypotheses according to the attested word forms of the hypothetical inflectional paradigm in the corpus

The system's general architecture is depicted in figure 1. The hypothesis generation step is explained in more detail in subsection 2.2.. The hypothesis selection step is explained in more detail in subsection 2.3..

2.2. Hypothesis Generation

Both the lexical module (called 'SLEX') and the hypothesis generation module ('SLES') are finite-state transducers which are based on the inflectional component of SMOR, a computational model of German morphology (Schmid et al. 2004; the inflectional component goes back to Schiller 1996). Just as SMOR, they are implemented with SFST, a free software package that provides a programming language and the necessary tools for compiling and running large finite-state transducers (Schmid, 2005), featuring two-level rules (Koskeniemmi, 1984) as well as cascaded architectures using transducer composition as proposed in (Kaplan & Kay, 1994). The advantages of using finite-state transducers for this lexicon acquisition task are that finite-state transducers are a wide-spread and well-established device for describing morphological patterns declaritively, that the same rule base, if carefully designed, can be used to support both analysis and generation by just switching the direction of the transducer, that the necessary development tools are readily available and are efficient enough to process large amounts of data, and that the formalism makes it very easy to reuse other transducers by using concatenation and composition.

SMOR and SLES have in common that they concatenate – at least at some level of the analysis – lexical entries, which consist of an inflectional stem (e.g., *Reim*, 'rhyme') and a symbol for an inflectional class (e.g., <NMasc_es_e>)[2], with all possible affixes for the specific inflectional class. Both sets – inflectional stems and mappings from inflectional classes to the set of possible inflectional markers – are represented as transducers. This concatenated transducer is then composed with an additional transducer from SMOR, applying morphophonological and orthographic rules, which are triggered by special symbols (for example, <UL> for 'umlaut' or <^Del> an e-elision in certain inflectional classes) in the set of inflectional markers.

The resulting transducer maps word forms to their underlying lexical representations, and vice versa. Whereas the lexical representations of a word form in SMOR comprise the lemma (citation form, e.g. the infinitive form for verbs) and the morphosyntactic properties, the lexical representation in SLES consist of the lexical entry and (optionally) the morphosyntactic properties. SLES can thus be used to determine the set of all possible lexical entries that could give rise to a given word form. This is achieved by using generic inflectional stems (basically .*) for each inflectional class[3]. The SLES transducer is designed to be run in both directions. It can be used to determine the set of possible lexical entries of a word form but it can as well be used to generate the whole set of predicted word forms – that is, the inflectional paradigm – of a lexical entry. An example of running SLES on a given word form, and on a given (hypothetical) lexical entry, is given in figure 2.

Once a lexical entry hypothesis is accepted, it can be entered into a lexicon, called SLEX. The SLEX transducer combines the functionality of SLES with the lexicon, by composing the lexical representation of SLES with the disjunction of all entries of a plain text file containing approved lexical entries. Thus, SLEX has basically the same capabilities as the SLES transducer, with the only difference that it outputs confirmed rather than hypothetical lexical entries. SLEX is also used to classify word forms into known and unknown forms. For the experiments reported here, SLEX contained lexical entries for roughly 500 lexemes[4] for closed word classes such as adpositions (112 lexemes), pro-forms (104), irregular verbs (196), irregular ad-

[2]SFST supports multi-character symbols, which are enclosed in angle brackets.

[3]It was necessary to constrain the form of stems for avoiding both spurious hypotheses during hypothesis generation as well as unwanted interactions between the morpho-phonology transducer when generating word forms from a (hypothetical) lexical representation. For example, inflectional stems of nouns with an inflectional class where most forms are formed by attaching *-e*, *-es*, *-en* should not end with *e*. Thus, the set of permissible stems is defined as a disjunction of patterns as the following:

```
[#LETTER#][#letter#]+[^e][<NMasc_es_e><NNeut_es_e>]
```

[4]Note that irregular, particularly suppletive, paradigms are often represented by several lexical entries.

```
analyze> Reime
Reim<NMasc_es_e>
Reime<NFem_0_n>
Reime<NNeut_s_s>
Reime<NNeut_s_0>
Reime<NMasc_s_s>
Reime<NFem_0_s>
Reime<NNeut_s_n>
Reime<NMasc_s_0>
Reime<NMasc_n_n>
Reime<NMasc_s_n>
Reim<NNeut_es_e>
```

```
generate> Reim<NMasc_es_e>
Reim
Reims
Reime
Reimes
Reimen
generate> Reime<NFem_0_s>
Reime
Reimes
generate> Reime<NNeut_s_s>
Reime
Reimes
```

Figure 2: Mapping word forms to lexical entry hypotheses, and vice versa

jectives (19) and miscellaneous other classes (61).

SLES basically uses the inflectional classes from SMOR, with some minor adaptions where a tighter control over the set of word forms was needed, for example, for having a single lexical representation for lexemes comprising several lexical entries as for irregular verbs. Unlike SMOR, SLES in general does not decompose inflectional stems using a proper word formation grammar. Conversion, in particular, such as from a past participle form of a verb to an adjective, is not permitted by the inflectional classes.

Particle verbs present a challenge for the acquisition system as outlined so far. Particle verbs[5] are combinations of a "verbal particle" with a verb, which usually also occurs as a lexeme of its own. They are special in that they are written as one orthographic word in subordinate clauses but written separately in main clauses. Furthermore, they incorporate an affix in two cells of the inflectional paradigm, namely the infinitive with incorporated *-zu-* and the past participle with incorporated *-ge-*. This is problematic since the position where *-zu-* and *-ge-* are to be inserted is not predictable if the token currently being analysed doesn't exhibit an incorporated *-zu-* or *-ge-*. Therefore, all possible positions for these infixes have to be considered. This increases the number of possible hypotheses dramatically[6], and unless no form with incorporated *-zu-* or *-ge-* is attested, it is not possible without further means to disambiguate the hypothesised analyses. Particle verbs are a mild exception to the rule that SLES does not perform a word-formation analysis since the irregular verbs taken from SLEX may function as the head of an inflectional stem for particle verbs. In this case, the position of the incorporated *-zu-* and *-ge-* is immediatly to the left of the head.

[5]For a discussion of the status of 'particle verbs' within a theory of grammar see Lüdeling (2001).

[6]Here are some examples for *einleuchten*, '(to) become clear', with the particle *ein*, 'into', and the verb *leuchten*, '(to) shine', with the infinitive *einzuleuchten* with incorporated *-zu-* and the past participle *eingeleuchtet*:

```
<ge>einleucht<VVReg> ei<ge/zu>nleuch<VVReg>
ei<ge/zu>nleucht<VVReg> ein<ge/zu>leuch<VVReg>
ein<ge/zu>leucht<VVReg> einl<ge/zu>euch<VVReg>
einl<ge/zu>eucht<VVReg> einle<ge/zu>uch<VVReg>
einle<ge/zu>ucht<VVReg> ...
```

2.3. Hypothesis Selection

The generation of lexical entry hypotheses never leads to a unique hypothesis for a particular word form because virtually all inflectional classes license at least one word form with a zero-suffix, as for example, the nominative singular of nouns, or the imperative singular of verbs. This means that there will be several lexical entry hypotheses for each word form with the very word form as the inflectional stem. For example, the nominal word form *Reim*, 'rhyme', can be the nominative singular of any inflectional class for nouns that is included in the system. Similarly, the word form *gereimt*, past participle of the verb REIMEN, '(to) rhyme' is not only analysed as a possible word form with the verbal stem *reim-* but also as a possible word form of hypothetical a verbal stem *gereim-* (similar to *gehör-*, '(to) belong to') or as an adjectival stem, which is in fact also correct.

Due to the over-generation of lexical entry hypotheses in the first step, the correct hypotheses have to be selected in a second step. Hypothesis selection takes places after all hypotheses for all word forms of the corpus have been determined. At this stage, it is possible to relate the hypotheses to the whole corpus. We can now ascertain how many word form types and how many tokens per word form type, gave rise to a particular hypothesis. This information can be used for establishing a score of the plausibility of a hypothesis.

How can we operationalise the plausibility of a lexical entry hypothesis? A straightforward idea would be to prefer one lexical entry hypothesis over another if it were attested by more different word forms. This is what I call the Majority Vote score. Assume for example that the word forms *Reim* and *Reime* are attested in a corpus. Using the Majority Vote criterion, the lexical entry hypothesis `Reim<NMasc_es_e>` in figure 2 would then win over `Reime<NFem_0_s>` and `Reime<NNeut_s_s>` since more word forms are attributed to this lexical entry.

The Majority Vote score favours lexical entries with large sets of different attested word forms. It does not consider the set of predicted word forms at all. If several hypotheses cover the same word forms then it seems reasonable to choose the hypotheses that make the fewest assumptions. This is achieved by determining the proportion of attested word forms among the predicted ones. This is what I call the Paradigm Attestation score.

The Paradigm Attestation score is not free of problems as well. Using this criterion, any hypothesis with fewer predicted word forms will be preferred to a hypothesis with more predicted but less attested word forms (compare for example Reim<NMasc_es_e> with Reim<NMasc_es_e> in figure 2). Since there are many inflectional classes that generate only two different word forms – for instance one with a zero-suffix, another with an *s*-suffix – these classes will usually be preferred by the Paradigm Attestation score. As a solution, we have to trade off both scores so that both the paradigm attestation as well as the number of attested word forms are taken into account. This is achieved by taking the product of the paradigm attestation proportion with the majority vote. This score is called Majority Vote $\times$ Paradigm Attestation (MVxPA).

In order to create a rank of all hypotheses, the following steps are applied: First, all competing hypotheses are determined. Competing hypotheses are groups of hypotheses that i) have the same part-of-speech, and ii) share attested word forms.[7] Hypotheses that generate exactly the same set of word forms (which is systematically the case for certain masculine and neuter noun classes) are conflated. For each such group, the hypotheses are sorted by their score as the first sort key, and by the token frequency attributed to each hypothesis as the second sort key. This establishes a local rank of competing hypotheses. These hypotheses can then be presented to a human annotator who manually checks the proposed lexical entry candidates for correctness. To this end, the groups of competing hypotheses should finally be ranked by the attributed token frequency of their highest ranked member, the rationale being that the most important lexical entry candidates in terms of coverage should be listed first.

3. Evaluation

The best way to evaluate the list of proposed lexical entries would be to compare it to an authorative list of correct lexical entries for a given corpus. Such a list, however, is not available. Instead of creating such a list for a given corpus by hand, I have opted for a different approach, namely, to create a gold standard by combining several other existing gold standard resources.

I have used the TIGER Corpus (Brants et al., 2002) as a learning corpus as it provides gold annotation for lemmas and morphosyntactic properties. Furthermore, I was able to use SMOR with the complete lexicon for the evaluation of the system. A direct comparison of the lexical data of SMOR and the proposed lexical entries of SLES is not always possible, though. Many lexical words are not contained as a separate lexical entry within SMOR's lexicon but are rather composed from smaller lexical items by productive word-formation patterns. I have therefore modified SMOR in such a way that it returns the set of all lexical entries for a word form in the format that is expected by SLES, that is, special symbols that mark morpheme boundaries are removed and the inflectional stem and the inflectional class are reconstructed. This version of SMOR has been run against the word forms in TIGER. In order to exclude remaining ambiguities that might arise from different or wrong word formation analyses, I have matched the list of lexical entries obtained by the modified SMOR against the annotations in TIGER. The basic assumption was that a lexical entry obtained by running SMOR on TIGER was only valid if it was compatible with the lemma, part-of-speech, and inflectional properties that were annotated in TIGER.

The aim of the evaluation is to assess i) the quality of the 'local' ranking of competing lexical entry hypotheses, and to show ii) how successful the method is in establishing a 'global' ranking method. This should be directly relevant to a human annotator: ideally, the precision of the proposed entries starts at a high value and decreases with increasing rank. If the precision drops beyond a certain value, the acquisition process can be stopped.

For the computation of precision, each position on the list of ranked hypotheses is counted as correct if any of the selected competing candidates can either be found in the SMOR lexicon or in the list of lexical entries obtained by running SMOR on TIGER. For recall, only the latter resource could be used as a reference since the targeted entries are those that underlie the word forms of the corpus; the aim is not to acquire the whole lexicon of the language, which is theoretically infinitely large.

Three scenarios are evaluated: in the first, only those competing lexical entry candidates are shown to the annotator that have the best score. In the case that several candidates have the same best score, all of them are assumed to be shown to the human annotator.[8] In the second and third scenarios, the number of included candidates is increased accordingly, using the same criterion as before. Here we look at all competing lexical entry candidates with the two or three best scores, respectively. The distribution of the number of competing hypotheses after selecting either the ones with the best score, the 2 best scores, or the 3 best scores is shown in figure 4.

The results of the evaluation are shown in figure 5. The statistical characteristics of the TIGER corpus is shown in figure 3. These first results look quite promising. At least for the first 25% of the ranked list, every second suggestion made by the system based on the lexical entry candidates with the highest score is correct. In this case, the annotator will be confronted with usually only one lexical entry hypothesis and asked for a judgement. A higher number of proposed candidates yields better results, especially for verbs where more than 80% precision are achieved.

4. Discussion

The present work demonstrates how knowledge about inflectional paradigms alone can be utilised to acquire lexical entries for an inflectional lexicon from a raw corpus. The acquisition process is based only on the tokens and their

[7]An alternative approach would be to first select hypotheses with high scores among the list of all hypotheses, and then remove all lexical entries on the list of hypotheses that were outranked by a higher ranked lexical entry.

[8]This is also true for lexical entry hypotheses that are formally equivalent. Since they predict exactly the same set of word forms, between which the proposed method cannot decide, these hypotheses have been conflated before.

frequencies in a corpus but does not require any further information. Thus, it is well suited for acquiring a candidate list of lexical entries that can be manually checked much more efficiently as opposed to manually creating such a list of lexical entries.
A flaw of the current system is that it is not robust to sparse data and errors in the corpus. One possible way to make the system more robust and to improve precision would be to use a probabilistic model over inflectional slots to estimate the plausibility of a word form. This, however, requires the presence of morphologically annotated corpora.
Another natural extension of the current system would be to exploit (categorial) agreement and government relations in the syntactic contexts of word forms, in addition to formal differences between word forms, as it is done in the 'Durm lemmatiser' for German nouns (Perera & Witte, 2005). Such a syntactic approach is orthogonal to the one presented here.

Acknowledgements

This work was mainly carried out at the Humboldt-Universität zu Berlin, as part of my diploma thesis under the supervision of Prof. Anke Lüdeling and Prof. Ulf Leser. I wish to express my sincere thanks to Stefanie Dipper and Viktor Trón for fruitful discussions on the topics covered here, and to Stefan Evert and Marco Baroni who helped me with the development of the original idea. I am very grateful to the IMS Stuttgart, and, in particular, to Uli Heid, for granting me access to the full SMOR lexicon, as well as to Helmut Schmid for the release of SFST and SMOR as free software. Finally, I want to thank the DFKI for supporting my participation at the LREC conference.

References

Brants, S., Dipper, S., Hansen, S., Lezius, W. & Smith, G. (2002). The TIGER Treebank. In *Proceedings of the Workshop on Treebanks and Linguistic Theories (TLT 2002)*. Sozopol, Bulgaria.

Clément, L., Sagot, B. & Lang, B. (2004). Morphology based automatic acquisition of large-coverage lexica. In *Proceedings of the 4th International Conference of Language Resources and Evaluation (LREC 2004)*. Lisbon, Portugal.

Forsberg, M., Hammarström, H. & Ranta, A. (2006). Morphological Lexicon Extraction from Raw Text Data. In *Proceedings of FinTAL - 5th International Conference on Natural Language Processing*. Turku, Finland.

Kaplan, R. M. & Kay, M. (1994). Regular Models of Phonological Rule Systems. *Computational Linguistics*, 20(3), pp. 331–378.

Koskeniemmi, K. (1984). A general computational model for word-form recognition and production. In *Proceedings of the 10th International Conference on Computational Linguistics (COLING 84)*, pp. 178–181. Stanford University, California, USA.

Lezius, W. (2000). Morphy - German Morphology, Part-of-speech Tagging and Applications. In *Proceedings of EURALEX 2000*. Stuttgart, Germany.

Lüdeling, A. (2001). *On Particle Verbs and Similar Constructions in German*. Stanford, California, USA: CSLI Publications.

Oliver, A., Castellón, I. & Màrquez, L. (2003). Automatic Lexical Acquisition from Raw Corpora: An Application to Russian. In *Proceedings of the EACL-2003 Workshop on Morphological Processing of Slavic Languages*. Budapest, Hungary.

Oliver, A. & Tadić, M. (2004). Enlarging the Croatian Morphological Lexicon by Automatic Lexical Acquisition from Raw Corpora. In *Proceedings of the 4th International Conference of Language Resources and Evaluation (LREC 2004)*, pp. 1259–1262. Lisbon, Portugal.

Perera, P. & Witte, R. (2005). A Self-Learning Context-Aware Lemmatizer for German. In *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP 2005)*, pp. 636–643. Vancouver, Canada.

Sagot, B. (2005). Automatic acquisition of a Slovak Lexicon from a Raw Corpus. In Matoušek, V., Mautner, P. & Pavelka, T. (Eds.), *Text, Speech and Dialogue: 8th International Conference, TSD 2005, Karlovy Vary, Czech Republic, September 12-15, 2005. Proceedings.*, vol. 3658 of *Lecture Notes in Computer Science*, pp. 156–163. Berlin / Heidelberg: Springer.

Schiller, A. (1996). Deutsche Flexions- und Kompositionsmorphologie mit PC-KIMMO. In Hausser, R. (Ed.), *Linguistische Verifikation. Dokumentation zur Ersten Morpholympics 1994*. Tübingen: Max Niemeyer Verlag.

Schmid, H. (2005). SFST Manual. Part of the SFST software package.

Schmid, H., Fitschen, A. & Heid, U. (2004). SMOR: A German Computational Morphology Covering Derivation, Composition, and Inflection. In *Proceedings of the 4th International Conference of Language Resources and Evaluation (LREC 2004)*, pp. 1263–1266. Lisbon, Portugal.

	tokens	types
word forms in TIGER	$888,299$	$89,385$
punctuation	$119,554$	23
word form hapax legomena	$45,224$	$45,224$
lemma hapax legomena	$30,143$	$30,143$
word forms used in acquisition	$740,846$	$79,249$
coverage initial SLEX transducer	$316,675$	$1,474$
coverage all hyps with the best score	$167,495$	$17,553$
coverage all hyps with the 2 best scores	$192,695$	$22,727$
coverage all hyps with the 2 best scores	$198,830$	$23,824$

Figure 3: Corpus statistics

nr.	best score	2 best scores	3 best scores
1	$13,494$	$1,035$	$1,035$
2	$3,563$	$8,778$	$1,959$
3	53	$6,407$	$5,150$
4	12	761	$5,112$
5		63	$1,454$
6		74	565
7		4	522
8			464
9			271
			$\ldots$

Figure 4: Distribution for the number of competing hypotheses after selecting either the ones with the best score, the 2 best scores, or the 3 best scores. That is, if among all sets of competing hypotheses only the hypotheses with the highest scores are retained, we find e.g. $3,563$ hypotheses that still competes with another hypothesis in the same group.

	rank up to	all	nouns	adjectives	verbs
hyps by the best score	25%	49.8% / 18.7%	48.1% / 15.0%	52.3% / 19.5%	74.5% / 37.5%
hyps by the best score	50%	49.3% / 37.8%	44.9% / 29.0%	44.0% / 33.3%	52.8% / 54.1%
hyps by the best score	75%	42.8% / 49.2%	43.7% / 42.9%	32.5% / 37.6%	39.4% / 59.1%
hyps by the best score	100%	37.7% / 57.8%	41.7% / 55.0%	32.1% / 49.0%	37.8% / 75.8%
hyps by the 2 best scores	25%	61.1% / 22.9%	59.3% / 18.6%	70.9% / 26.9%	83.1% / 41.0%
hyps by the 2 best scores	50%	64.1% / 49.3%	57.4% / 37.0%	55.9% / 42.8%	59.5% / 59.5%
hyps by the 2 best scores	75%	55.9% / 64.0%	57.0% / 55.8%	51.9% / 60.5%	47.6% / 69.1%
hyps by the 2 best scores	100%	51.0% / 77.9%	58.2% / 76.6%	46.7% / 72.4%	45.2% / 88.0%
hyps by the 3 best scores	25%	67.3% / 25.4%	68.9% / 21.7%	71.6% / 27.1%	84.1% / 41.5%
hyps by the 3 best scores	50%	68.3% / 52.6%	64.6% / 41.8%	56.3% / 43.1%	60.6% / 60.8%
hyps by the 3 best scores	75%	59.2% / 67.8%	63.1% / 61.9%	52.2% / 60.8%	49.7% / 72.2%
hyps by the 3 best scores	100%	53.7% / 82.2%	63.1% / 83.2%	46.9% / 72.8%	47.5% / 92.8%

Figure 5: Precision and recall for the Majority Vote $\times$ Paradigm Attestation rank.