Cross-Language Text Classification using Structural Correspondence Learning

Peter Prettenhofer and Benno Stein

Web Technology & Information Systems Group
Bauhaus-Universität Weimar

September 15, 2010
Outline

Cross-Language Text Classification

Cross-Language Structural Correspondence Learning

Empirical Results
Outline

Cross-Language Text Classification

Cross-Language Structural Correspondence Learning

Empirical Results
Problem Statement
Create a classifier for a text classification task in some target language $\mathcal{T}$ given labeled examples for the identical task in a different source language $\mathcal{S}$.

- Example: Create a sentiment classifier for German book reviews given training book reviews written in English.
- Can be cast as a domain adaptation problem.
Text Classification

- We assume BoW document representations $x$ and linear classifiers $w$.
- For simplicity, we consider binary classification, $y \in \{-1, +1\}$.

```
| Term   | x  | w  |
|--------|----|----|
| great  | 0.5| 2.1|
| book   | 0.2| -0.2|
| money  | 0  | -1.9|
| ...    | ...| ...|
| movie  | 0  | -0.6|
| poetry | 0.3| 0.9 |
| pages  | 0  | -1.1|
```

- Training: infer $w$ from a set of training examples $D_S = \{(x_i, y_i)\}$. 

$\text{sign}(w^T x)$
Disjoint vocabulary

- Vocabulary divides into $V_S$ and $V_T$ with $V_S \cap V_T = \emptyset$.
- A linear classifier trained on $D_S$ can associate non-zero weights only with $V_S$.
Cross-Language Text Classification (2)

Cross-lingual representation

- A concept space that underlies both languages.
- Let $\theta$ denote a (linear) map from the original to the cross-lingual representation.

$$\theta x \rightarrow \text{sign}(v^T \theta x)$$
Cross-Language Text Classification (3)

- $\theta$ encodes cross-lingual word correspondences.

- Current approaches use various linguistic resources to construct $\theta$:
  - Bilingual dictionary.
  - Parallel corpus.
  - Machine translation (MT) system.
Cross-Language Text Classification (3)

- $\theta$ encodes cross-lingual word correspondences.

- Current approaches use various linguistic resources to construct $\theta$:
  - Bilingual dictionary.
  - Parallel corpus.
  - Machine translation (MT) system.

- Our approach learns $\theta$ from unlabeled data.
Outline

Cross-Language Text Classification

Cross-Language Structural Correspondence Learning

Empirical Results
Cross-Language Structural Correspondence Learning

- CL-SCL uses unlabeled data and a word translation oracle to induce cross-lingual word correspondences.
- Builds on Structural Correspondence Learning (SCL) [Blitzer et al, 2006].
- Advantages:
  - Task specific correspondences.
  - Efficiency in terms of linguistic resources.
  - Efficiency in terms of computational resources.
- Competitive or better than MT while requiring fewer resources.
CL-SCL uses unlabeled data and a word translation oracle to induce cross-lingual word correspondences.

Builds on Structural Correspondence Learning (SCL) [Blitzer et al, 2006].

Advantages:

- Task specific correspondences.
- Efficiency in terms of linguistic resources.
- Efficiency in terms of computational resources.

Competitive or better than MT while requiring fewer resources.
CL-SCL - Learning Setting

1. **Labeled source data** $D_S$.
2. **Unlabeled data** $D_u = D_{S,u} \cup D_{T,u}$
3. **Translation oracle** $o : V_S \rightarrow V_T$

```
\begin{align*}
\text{Words in } V_S & \quad \text{Words in } V_T \\
x = (x_1, \ldots, x_{|V|}) & \quad \ldots, x_{|V|}
\end{align*}
```

- **Term frequencies**
- **Positive class label**
- **Negative class label**
- **No value**
Step 1 - Pivot Selection

- A **pivot** is a pair of words \( \{w_S, w_T\} \).
- Pivots have to satisfy the following conditions:
  - **Confidence**: Both words are correlated with the class label.
  - **Support**: Both words occur frequently in \( D_{S,u} \) and \( D_{T,u} \).
- Example: \( \{\text{excellent}_S, \text{exzellent}_T\} \).
Step 1 - Pivot Selection

- A pivot is a pair of words $\{w_S, w_T\}$.
- Pivots have to satisfy the following conditions:
  - **Confidence**: Both words are correlated with the class label.
  - **Support**: Both words occur frequently in $D_{S,u}$ and $D_{T,u}$.
- Example: $\{\text{excellent}_S, \text{exzellent}_T\}$.

**Heuristic**

1. Select subset from $V_S$ according to MI w.r.t. $D_S$.
2. Translate words into $T$.
3. Eliminate pivots which occur less than $\phi$ times in $D_u$. 
Step 1 - Pivot Selection

- A pivot is a pair of words \( \{w_S, w_T\} \).
- Pivots have to satisfy the following conditions:
  - **Confidence:** Both words are correlated with the class label.
  - **Support:** Both words occur frequently in \( D_{S,u} \) and \( D_{T,u} \).
- Example: \( \{\text{excellent}_S, \text{exzellent}_T\} \).

**Heuristic**

1. Select subset from \( V_S \) according to MI w.r.t. \( D_S \).
2. Translate words into \( T \).
3. Eliminate pivots which occur less than \( \phi \) times in \( D_u \).

Let \( m \) denote the number of pivots.
Step 2 - Train Pivot Classifiers (1)

- Model the correlations between each pivot and all other words.
- **Pivot classifier**: A linear classifier that predicts whether or not $w_S$ or $w_T$ occur in a document.
Step 2 - Train Pivot Classifiers (2)

- Let $w_l$ denote the pivot classifier for the $l$-th pivot $\{w_S, w_T\}$.
- $w_l$ captures both the correlation between $w_S$ and $V_S \setminus w_S$ and between $w_T$ and $V_T \setminus w_T$.
  - Implicitly aligns non-pivot words from both $V_S$ and $V_T$. 
Step 2 - Train Pivot Classifiers (2)

▶ Let $w_l$ denote the pivot classifier for the $l$-th pivot \{$w_S, w_T$\}.
▶ $w_l$ captures both the correlation between $w_S$ and $V_S \setminus w_S$ and between $w_T$ and $V_T \setminus w_T$.
   ▶ Implicitly aligns non-pivot words from both $V_S$ and $V_T$.

Example: \{$boring_S$, langweilig$_T$\}

langatmig (lengthy), spannung (tension), war (was), characters, handlung (story), pages, finish, seiten (pages), story
Step 3 - Compute SVD

- If two words (e.g., pages\(_S\) and seiten\(_T\)) are correlated across a number of pivots we assume correspondence between them.

- Identify correlations across pivots by computing the SVD of the parameter matrix \(W\),

\[ W = [w_1 \cdots w_m] \]

- Let \(\theta^T\) be the top-\(k\) left singular vectors of \(W\).

- At training and test time simply apply \(\theta x\) for each instance \(x\).
Alternative View

- Use $\theta$ to constraint the parameter space for the target task [Ando & Zhang, 2005].
Alternative View

- Use $\theta$ to constraint the parameter space for the target task [Ando & Zhang, 2005].
Alternative View

- Use $\theta$ to constraint the parameter space for the target task [Ando & Zhang, 2005].

\[
\mathbf{w} = \theta^T \mathbf{v}
\]
Computational Considerations

- SVD is the computational bottleneck if $W$ is large.

- Make $W$ sparse:
  - Set negative values to zero [Ando & Zhang, 2005; Blitzer et al., 2007; Prettenhofer & Stein, 2010a].
  - Use sparse regularization for pivot classifiers [Prettenhofer & Stein, 2010b].

Elastic-Net Regularization [Zou & Hastie, 2005]
- A convex combination of L2 and L1 norm penalties, $R(w) = \alpha \|w\|_2^2 + (1 - \alpha) \|w\|_1$.
- Superior to L1 penalty when handling highly correlated features.
Computational Considerations

- SVD is the computational bottleneck if $\mathbf{W}$ is large.
- Make $\mathbf{W}$ sparse:
  - Set negative values to zero [Ando & Zhang, 2005; Blitzer et al, 2007; Prettenhofer & Stein, 2010a].

Elastic-Net Regularization [Zou & Hastie, 2005]

- A convex combination of L2 and L1 norm penalties, $R(\mathbf{w}) = \alpha \|\mathbf{w}\|_2^2 + (1 - \alpha) \|\mathbf{w}\|_1$.
  - Superior to L1 penalty when handling highly correlated features.
Computational Considerations

- SVD is the computational bottleneck if $W$ is large.
- Make $W$ sparse:
  - Set negative values to zero [Ando & Zhang, 2005; Blitzer et al., 2007; Prettenhofer & Stein, 2010a].
  - Use **sparse regularization** for pivot classifiers [Prettenhofer & Stein, 2010b].
Computational Considerations

- SVD is the computational bottleneck if $W$ is large.
- Make $W$ sparse:
  - Set negative values to zero [Ando & Zhang, 2005; Blitzer et al., 2007; Prettenhofer & Stein, 2010a].
  - Use **sparse regularization** for pivot classifiers [Prettenhofer & Stein, 2010b].

**Elastic-Net Regularization [Zou & Hastie, 2005]**

- A convex combination of L2 and L1 norm penalties,

$$R(w) = \alpha \|w\|_2^2 + (1 - \alpha) \|w\|_1.$$  

- Superior to L1 penalty when handling highly correlated features.
Outline

Cross-Language Text Classification

Cross-Language Structural Correspondence Learning

Empirical Results
Experimental Setup (1)

Data: Amazon product reviews

- Categories: Books, dvd, and music.
- Source language: English.
- Target language: German, French, and Japanese.
- Nine $S-T$-category combinations.
  - 2,000 training and 2,000 test examples (balanced).
  - 10,000 - 50,000 unlabeled examples from each language.

Training via Stochastic Gradient Descent

- Smoothed hinge loss as loss function.
- L2 penalty for target task.
- Elastic-Net for pivot classifiers.

Fast: 2-10sec / pivot classifier.
Experimental Setup (1)

Data: Amazon product reviews

- Categories: Books, dvd, and music.
- Source language: English.
- Target language: German, French, and Japanese.
- Nine $S$-$T$-category combinations.
  - 2,000 training and 2,000 test examples (balanced).
  - 10,000 - 50,000 unlabeled examples from each language.

Training via Stochastic Gradient Descent

- Smoothed hinge loss as loss function.
- L2 penalty for target task.
- Elastic-Net for pivot classifiers.
- Fast: 2-10sec / pivot classifier.
Experimental Setup (2)

- **Upper Bound (UB):**
  - Classification performance if training data in $T$ is available.

- **Baseline (CL-MT):**
  - Translate test documents into $S$ with Google Translate.

- **CL-SCL:**
  - Uses 450 pivots, dimensionality reduction to $k = 100$, $|D_u| \approx 10^5$, and $\alpha = 0.85$.
  - Google Translate as translation oracle.
### Results

| \( \mathcal{T} \) | Cat. | \( \mu \) | \( \mu \) | \( \Delta \) | \( \mu \) | \( \Delta \) | RR[\%] |
|------------------|------|------|------|------|------|------|------|
|                  | books | 83.79 | 79.68 | 4.11 | † 83.34 | 0.45 | 89.05% |
| German           | dvd   | 81.78 | 77.92 | 3.86 | † 80.89 | 0.89 | 76.94% |
|                  | music | 82.80 | 77.22 | 5.58 | † 82.90 | -0.10 | 101.79% |
|                  | books | 83.92 | 80.76 | 3.16 |       |      |       |
| French           | dvd   | 83.40 | 78.83 | 4.57 |       |      |       |
|                  | music | 86.09 | 75.78 | 10.31 |       |      |       |
|                  | books | 78.09 | 70.22 | 7.87 | †† 77.00 | 1.09 | 86.15% |
| Japanese         | dvd   | 81.56 | 71.30 | 10.26 | †† 76.37 | 5.19 | 49.42% |
|                  | music | 82.33 | 72.02 | 10.31 | †† 77.34 | 4.99 | 51.60% |

- ~ 60% reduction in relative error due to cross-lingual adaptation.
| $\mathcal{T}$ | Cat.    | UB $\mu$ | CL-MT $\mu$ | $\Delta$ | CL-SCL $\mu$ | $\Delta$ | RR[\%] |
|------------|---------|----------|-------------|----------|-------------|----------|--------|
| German     | books   | 83.79    | 79.68       | 4.11     | 83.34       | 0.45     | 89.05% |
|            | dvd     | 81.78    | 77.92       | 3.86     | 80.89       | 0.89     | 76.94% |
|            | music   | 82.80    | 77.22       | 5.58     | 82.90       | -0.10    | 101.79% |
| French     | books   | 83.92    | 80.76       | 3.16     | 81.27       | 2.65     | 16.14% |
|            | dvd     | 83.40    | 78.83       | 4.57     | 80.43       | 2.97     | 35.01% |
|            | music   | 86.09    | 75.78       | 10.31    | 78.05       | 8.04     | 22.02% |
| Japanese   | books   | 78.09    | 70.22       | 7.87     | 77.00       | 1.09     | 86.15% |
|            | dvd     | 81.56    | 71.30       | 10.26    | 76.37       | 5.19     | 49.42% |
|            | music   | 82.33    | 72.02       | 10.31    | 77.34       | 4.99     | 51.60% |

▶ ~ 60% reduction in relative error due to cross-lingual adaptation.
Results

| $\mathcal{T}$ | Cat. | UB | CL-MT | CL-SCL | RR[\%] |
|---------------|------|----|-------|--------|--------|
|               |      | $\mu$ | $\mu$ | $\Delta$ | $\mu$ | $\Delta$ |
| books | 83.79 | 79.68 | 4.11 | † | 83.34 | 0.45 | 89.05% |
| German | dvd | 81.78 | 77.92 | 3.86 | † | 80.89 | 0.89 | 76.94% |
| | music | 82.80 | 77.22 | 5.58 | † | 82.90 | -0.10 | 101.79% |
| books | 83.92 | 80.76 | 3.16 | | 81.27 | 2.65 | 16.14% |
| French | dvd | 83.40 | 78.83 | 4.57 | | 80.43 | 2.97 | 35.01% |
| | music | 86.09 | 75.78 | 10.31 | | 78.05 | 8.04 | 22.02% |
| books | 78.09 | 70.22 | 7.87 | †† | 77.00 | 1.09 | 86.15% |
| Japanese | dvd | 81.56 | 71.30 | 10.26 | †† | 76.37 | 5.19 | 49.42% |
| | music | 82.33 | 72.02 | 10.31 | †† | 77.34 | 4.99 | 51.60% |

▶ ~ 60% reduction in relative error due to cross-lingual adaptation.
## Results

| \( \mathcal{T} \) | Cat.   | UB  | CL-MT | CL-SCL | RR[\%] |
|------------------|--------|-----|-------|--------|--------|
|                  | \( \mu \) | \( \mu \) | \( \Delta \) | \( \mu \) | \( \Delta \) |
| books            | 83.79  | 79.68 | 4.11  | † 83.34 | 0.45  | 89.05% |
| German           |        |      |       |        |       |       |
| dvd              | 81.78  | 77.92 | 3.86  | † 80.89 | 0.89  | 76.94% |
| books            | 83.92  | 80.76 | 3.16  | †† 81.27 | 2.65  | 16.14% |
| French           |        |      |       |        |       |       |
| dvd              | 83.40  | 78.83 | 4.57  | † 80.43 | 2.97  | 35.01% |
| books            | 78.09  | 70.22 | 7.87  | †† 77.00 | 1.09  | 86.15% |
| Japanese         |        |      |       |        |       |       |
| dvd              | 81.56  | 71.30 | 10.26 | †† 76.37 | 5.19  | 49.42% |
| books            | 82.33  | 72.02 | 10.31 | †† 77.34 | 4.99  | 51.60% |
|                  |        |      |       |        |       |       |

\( \sim 60\% \) reduction in relative error due to cross-lingual adaptation.
| $\mathcal{T}$ | Cat. | UB | CL-MT | CL-SCL | RR [%] |
|---|---|---|---|---|---|
| | | $\mu$ | $\mu$ | $\Delta$ | $\mu$ | $\Delta$ | |
| books | 83.79 | 79.68 | 4.11 | † | 83.34 | 0.45 | 89.05% |
| German | dvd | 81.78 | 77.92 | 3.86 | † | 80.89 | 0.89 | 76.94% |
| | music | 82.80 | 77.22 | 5.58 | † | 82.90 | -0.10 | 101.79% |
| books | 83.92 | 80.76 | 3.16 | | 81.27 | 2.65 | 16.14% |
| French | dvd | 83.40 | 78.83 | 4.57 | | 80.43 | 2.97 | 35.01% |
| | music | 86.09 | 75.78 | 10.31 | | 78.05 | 8.04 | 22.02% |
| books | 78.09 | 70.22 | 7.87 | †† | 77.00 | 1.09 | 86.15% |
| Japanese | dvd | 81.56 | 71.30 | 10.26 | †† | 76.37 | 5.19 | 49.42% |
| | music | 82.33 | 72.02 | 10.31 | †† | 77.34 | 4.99 | 51.60% |

► ~ 60% reduction in relative error due to cross-lingual adaptation.
### Task-Specific Word Correlations

| Pivot | English | German |
|-------|---------|--------|
|       | Semantics | Pragmatics | Semantics | Pragmatics |
| \{beautiful$_S$, amazing, schön$_T$\} beautiful, beauty, lovely | picture, pattern, poetry, photographs, paintings | schöner, traurig | bilder, illustriert |
| \{boring$_S$, plain, langweilig$_T$\} asleep, dry, long | characters, pages, story | langatmig, einfach, enttäuscht | charaktere, handlung, seiten |

- Such task-specific correlations cannot be obtained from a general parallel corpus.
## Task-Specific Word Correlations

| Pivot       | English Semantics | English Pragmatics | German Semantics | German Pragmatics |
|-------------|-------------------|--------------------|------------------|-------------------|
| {beautiful$_S$, amazing, schön$_T$} | picture, pattern, poetry, photographs, paintings | schöner, traurig | bilder, illustriert |
| {boring$_S$, plain, langweilig$_T$} | characters, pages, story | langatmig, einfach, enttäuscht | charaktere, handlung, seiten |

▶ Such task-specific correlations cannot be obtained from a general parallel corpus.
The more unlabeled data the better.

Even a small number of pivots captures a large part of the correspondences between $S$ and $T$.

SVD is crucial to the success of CL-SCL.

- Value of $k$ is task-insensitive.
Cross-language text classification can be cast as a domain adaptation problem.

CL-SCL uses unlabeled data and a word translation oracle to induce task-specific, cross-lingual word correspondences.

Convincing empirical results.

Competitive or better than MT while requiring fewer resources.

Future work: apply CL-SCL to other NLP tasks.

E.g., cross-language named entity recognition.
Thanks! Questions?

Data: http://webis.de/research/corpora/
SGD-Code: http://github.org/pprett/bolt/
References

- Domain Adaptation using Structural Correspondence Learning
  [Blitzer, J., McDonald, R., and Pereira F., EMNLP, 2006]

- Domain Adaptation for Sentiment Classification
  [Blitzer, J., Dredze, M., and Pereira, F., ACL, 2007]

- A framework for learning predictive structures from multiple tasks and unlabeled data
  [Ando, R. K. and Zhang, T., JMLR, 2005]

- Regularization and variable selection via the elastic net
  [Zou, H. and Hastie, T., JRSS, 2005]

- Cross-Language Text Classification using Structural Correspondence Learning
  [Prettenhofer, P., and Stein, B., ACL, 2010a]

- Cross-Lingual Adaptation using Structural Correspondence Learning
  [Prettenhofer, P., and Stein, B., arXiv, 2010b]
Discriminative Training of Linear Classifiers

- Minimize the (regularized) training error,

\[
\arg \min_w \sum_{(x, y) \in D} L(y, w^T x) + \lambda R(w).
\]

- Loss term \( L \) measures model (mis)fit.
- Regularization term \( R \) penalizes model complexity.

\[
\begin{align*}
L2: \quad &R(w) = \|w\|_2^2 = \sum_i w_i^2 \\
L1: \quad &R(w) = \|w\|_1 = \sum_i |w_i| \\
Elastic-Net: \quad &R(w) = \alpha \|w\|_2^2 + (1 - \alpha) \|w\|_1
\end{align*}
\]
## Dataset Statistics

| $\mathcal{T}$ | Category | Unlabeled data | Labeled data | Vocabulary |
|---------------|----------|----------------|--------------|------------|
|               |          | $|D_{S,u}|$ | $|D_{T,u}|$ | $|D_S|$ | $|D_T|$ | $|V_S|$ | $|V_T|$ |
| German        | books    | 50,000         | 50,000       | 2,000     | 2,000     | 64,682 | 108,573 |
|               | dvd      | 30,000         | 50,000       | 2,000     | 2,000     | 52,822 | 103,862 |
|               | music    | 25,000         | 50,000       | 2,000     | 2,000     | 41,306 | 99,287  |
|               | books    | 50,000         | 32,000       | 2,000     | 2,000     | 64,682 | 55,016  |
| French        | dvd      | 30,000         | 9,000        | 2,000     | 2,000     | 52,822 | 29,519  |
|               | music    | 25,000         | 16,000       | 2,000     | 2,000     | 41,306 | 42,097  |
|               | books    | 50,000         | 50,000       | 2,000     | 2,000     | 64,682 | 52,311  |
| Japanese      | dvd      | 30,000         | 50,000       | 2,000     | 2,000     | 52,822 | 54,533  |
|               | music    | 25,000         | 50,000       | 2,000     | 2,000     | 41,306 | 54,463  |
|               | -        | 60,000         | 60,000       | 6,000     | 6,000     | 76,629 | 124,529 |
|               | -        | 60,000         | 45,000       | 6,000     | 6,000     | 76,629 | 74,807  |
|               | -        | 60,000         | 60,000       | 6,000     | 6,000     | 76,629 | 64,050  |
## Results

| $T$ | Cat.   | Upper Bound | CL-MT | CL-SCL |
|-----|--------|-------------|-------|--------|
|     |        | $\mu$       | $\sigma$ | $\mu$ | $\sigma$ | $\Delta$ | $\mu$ | $\sigma$ | $\Delta$ |
|     |        |             |       |       |         |         |       |       |         |
| books | 83.79 | ±0.20       |         | 79.68 | ±0.13       | 4.11     | † 83.34 | ±0.02        | 0.45     |
| German dvd | 81.78 | ±0.27       |         | 77.92 | ±0.25       | 3.86     | † 80.89 | ±0.02        | 0.89     |
| music | 82.80 | ±0.13       |         | 77.22 | ±0.23       | 5.58     | † 82.90 | ±0.00        | -0.10    |
| books | 83.92 | ±0.14       |         | 80.76 | ±0.34       | 3.16     | 81.27   | ±0.08        | 2.65     |
| French dvd | 83.40 | ±0.28       |         | 78.83 | ±0.19       | 4.57     | 80.43   | ±0.05        | 2.97     |
| music | 86.09 | ±0.13       |         | 75.78 | ±0.65       | 10.31    | 78.05   | ±0.06        | 8.04     |
| books | 78.09 | ±0.14       |         | 70.22 | ±0.27       | 7.87     | †† 77.00 | ±0.06        | 1.09     |
| Japanese dvd | 81.56 | ±0.28       |         | 71.30 | ±0.28       | 10.26    | †† 76.37 | ±0.05        | 5.19     |
| music | 82.33 | ±0.13       |         | 72.02 | ±0.29       | 10.31    | †† 77.34 | ±0.06        | 4.99     |

German - 92.95 ±0.11 92.25 ±0.07 0.70 92.61 ±0.06 0.34
French - 93.27 ±0.07 90.58 ±0.17 2.69 90.57 ±0.13 2.70
Japanese - 89.43 ±0.11 82.14 ±0.22 7.29 †† 85.03 ±0.10 4.40
## Effect of Regularization

| $\mathcal{T}$ | Category | $\mathcal{L}^2^+$ | L1 | Elastic-Net |
|---------------|----------|-------------------|----|-------------|
|               |          | $\mu$ | d[\%] | $\mu$ | d[\%] | $\mu$ | d[\%] |
| German        | books    | 79.50 | 17.88 | 82.45 | 1.24 | 83.34 | 11.02 |
|               | dvd      | 77.06 | 16.84 | 78.60 | 1.43 | 80.89 | 12.25 |
|               | music    | 77.60 | 16.00 | 81.41 | 1.72 | 82.90 | 13.92 |
|               | books    | 79.02 | 16.50 | 80.75 | 1.87 | 81.27 | 14.13 |
|               | dvd      | 78.80 | 19.23 | 78.70 | 3.98 | 80.43 | 23.22 |
|               | music    | 77.72 | 16.70 | 77.32 | 3.72 | 78.05 | 21.60 |
|               | books    | 73.09 | 15.21 | 71.06 | 1.27 | 77.00 | 10.47 |
| Japanese      | dvd      | 71.10 | 14.86 | 75.75 | 1.48 | 76.37 | 11.84 |
|               | music    | 75.15 | 13.72 | 76.22 | 1.83 | 77.34 | 13.39 |
| German        | -        | 89.69 | 16.19 | 88.73 | 0.92 | 92.61 | 8.38 |
| French        | -        | 87.59 | 16.29 | 89.65 | 1.36 | 90.57 | 11.37 |
| Japanese      | -        | 82.83 | 16.71 | 84.26 | 1.23 | 85.03 | 10.15 |