

Are We Making Ourselves Clear?

Terminology Management and Machine Translation at Volkswagen

Ulrike Bernardi, Andras Bocsak and Jörg Porsiel

FS-ZM on the behalf of Volkswagen Service Division
VK-3/1 Language Services
P.O. box 1997 D-38436 Wolfsburg
extern.joerg.porsiel@volkswagen.de

Comprendium Lingua GmbH Munich
Neumarkter Str. 81 D-81541 Munich
ulrike.bernardi@comprendium.de andras.bocsak@comprendium.de

Abstract. User acceptance and system integration in existing IT environments are the key factors in machine translation (MT) becoming a universally implemented tool. Although output quality obviously plays a major role, it is not the sole decisive factor. User perception of MT, however, is an important issue, and MT needs to be seen as one of several tools in the user's work environment. Trends, such as globalisation and the high volume of multi-lingual information available on the internet and within large companies, have made fast and consistent translation, and thus machine translation, indispensable in accelerating decision making. However, the logistics and workflows necessary for the successful integration of MT are complex and varied. Comprendium employs a highly flexible architecture to deal with these complex customer-specific demands. It thus advances the practical use of MT by allowing a wide range of MT workflow solutions to be implemented both in the translation "cottage industry" and in large corporate translation portals and professional translation environments. This paper gives an overview of two applications at Volkswagen in which the machine translation system has been integrated successfully with very good user acceptance. The applications are the "Volkswagen Language Portal" and the "Process Schedule" (*Arbeitsplan*).

1. Introduction

As is the case with other large international companies, Volkswagen needs to manage a range of issues arising from an increasing volume of information. The sheer volume of different types of information at Volkswagen Customer Service® clearly underlines the urgent need for a concerted and uniform approach. In addition to owner's manuals, approx. 40,000 pages of highly diverse publications are produced annually in the Workshop Information Department of Volkswagen's Headquarters in Wolfsburg.

Workshop manuals, current flow diagrams, maintenance tables, data sheets and software for diagnostic systems also contribute to this total.

As a global player, Volkswagen sells its vehicles in more than 150 countries, which means that information management tasks and associated potential problems can become very large. Owner's manuals, for example, are translated into twenty-seven languages, including English, French, Spanish, Chinese, Hebrew and Estonian. German is the source language of approximately 95% of all texts of this type. Just one small mistake in terminology or linguistic ambiguity in the German text can have repercussions in all the languages the text is translated into.

A sound and unified approach is needed to master challenges such as

- significant increase in technical documentation,
- greater vehicle complexity and associated technologies (e.g. electronics),
- drastically reduced innovation cycles,
- stricter EU legislation and other legal requirements,
- shorter time-to-market and,
- growing translation requirements.

Thus, approximately 2 years ago, Volkswagen initiated a terminology management project in its Service Division. The project was launched by Volkswagen's Language Services Department in Wolfsburg and involved translators, terminologists and technical authors from both Volkswagen and Audi.

The overall objective of the project was the assurance of high quality information and texts for internal and external use. Key elements are defined, standardised and consistent terminology. This verified terminology is to be used not only by translators and interpreters, but will also form the basis of a new editing tool to be implemented by the technical authors.

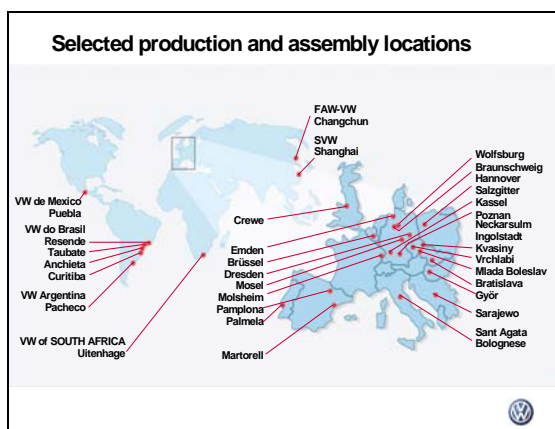


Figure 1. Selected VW production and assembly locations

2. Volkswagen Language Portal

2.1. Scenario

Some decisions can be based on the gist of a document, such as an e-mail, without having to know the exact contents of the document. Providing an exact understanding of a text is not crucial, MT can provide a gist translation in circumstances where there is no time to wait for the services of a human translator.

For company brochures, workshop information, and press and marketing information, top-quality translation is indispensable – and human translators will always be involved in translating texts of this nature. But even in a translator's environment, translation tools and automated processes assist in meeting deadlines, help reduce costs and relieve the load on internal and external translators.

Advanced, powerful and easy-to-use tools for internal communication within the Volkswagen Group were needed for gaining information quickly for business research, reporting and monitoring to overcome the language barriers between the company's 340,000+ employees in production, design, research and assembly locations all over the world. Machine translation is a tool which can help accelerate everyday processes by providing the user with "raw" or "gist translations" in a split-second for sources such as e-mails, reports, websites etc. thus helping the user gain a competitive edge when it comes to further action to be taken.

For standardisation and real-time availability of specific Volkswagen terminology, integration of terminological tools and large-scale terminology imports into the machine translation system proved to be indispensable. Thus, the basis for the Language Portal was formed, and Volkswagen terminology was integrated into MT engines to ensure high terminology consistency throughout the Group.

Providing terminology and machine translation services on a companywide, indeed a worldwide scale, is also a question of target group oriented supply of language-relevant applications. In an ideal world each target group would need and have its own linguists and terminologists. However, this is impractical, if not unmanageable. Nevertheless, Volkswagen strives to provide all internal customers (such as technical writers, translators, interpreters etc.) with the best services the Language Portal can offer in order to guarantee customer satisfaction inside as well as outside the company gates.

Volkswagen's terminology management project has three main components:

1 – A web-based **multilingual terminology database** containing some 12,000 revised terms in German, English, French, Italian, Swedish, and Spanish. In a few weeks a seventh language

will be added to the database, as the Czech car manufacturer Škoda (a member of the Volkswagen Group) recently joined the project.

Each of the entries comprises a definition, a context example, sources, comments and illustrations where necessary and helpful.

This customised multilingual dictionary will be made available via the Language Portal throughout the whole Group as a unique terminology and corporate language quality assurance system.

The terminology database is also vital for supplying verified terminology to the two following applications.

2 – **Machine translation** in German, English and Spanish for internal communication among Volkswagen Group employees.

3 – An **automated authoring system**: this complex tool supplies authors a range of functions such as checking of spelling, grammar, style, validated terminology and specific Volkswagen abbreviations. Terminology is imported into the system via an interface to the new terminology database which is maintained con-

tinuously by Volkswagen translators and terminologists.

2.2. Solution

The solution is the Volkswagen Language Portal on Volkswagen's own intranet, providing different applications, of which machine translation is one of the most frequently used.

Prior to the above-mentioned project, machine translation had already been in use at Volkswagen Coaching, which was responsible, among other things, for knowledge management at Volkswagen. However, after the terminology management project was launched, machine translation was handed over to the terminology project staff, as they proved to be the best qualified for this job.

After thorough internal testing over a long period and assessment the products of six major MT-providers regarding integration, performance, translation quality and availability, **Compendium's Translator Server** was chosen by Volkswagen Language Services. This service request distribution system provides translation

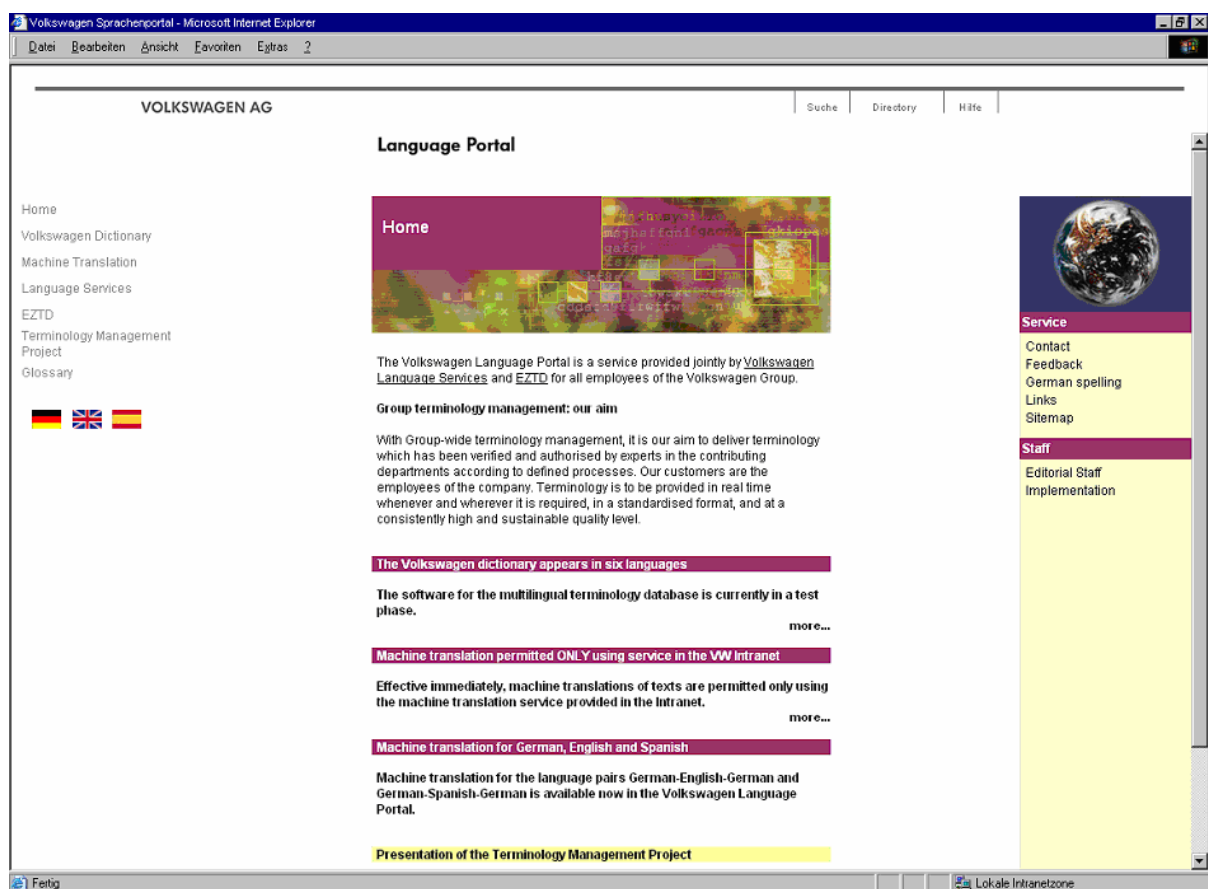


Figure 2. Volkswagen Language Portal in German, English and Spanish

services to clients, typically web clients or professional translator desktop clients. It can control operations in configurations involving large numbers of translation engines and web-clients (see section 4 for detailed description).

Translation directions are: German ⇔ English (English being the company's "lingua franca") and German ⇔ Spanish (Spanish being important to the Volkswagen Group because of SEAT based in Spain, and the production locations in Mexico and South America).

Compendium's monolingual and bi-lingual terminology extraction tools have also been applied to build the terminology base for the import into the MT lexica.

Volkswagen's Language Portal also contains a custom-made dictionary (based on TRADOS MultiTerm iX) with automotive-specific content in Volkswagen's six in-house languages. It is continuously enlarged, maintained and updated by terminologists and translators. This service is accessible worldwide in real-time on the Volkswagen intranet and will soon be enlarged in quality and quantity by Škoda's terminology.

2.3. Language Portal Statistics

The test phase of Volkswagen's Language Portal was launched in December 2002 together with a first test of the embedded machine translation software and ended in the summer of 2003. Sent tasks/hits totalled approximately 465,000 per month in a phase where access was limited to a very small number of specially selected users. Internal demand for fast gist translations turned out to be extremely high, much higher than expected.

General access to the Language Portal, and thus to Machine Translation, was provided on 21 September 2004: The number of hits within the first month increased by a stunning 2,198%.

3. Localizing Assembly Instructions at Volkswagen

3.1. Scenario

For a global player in the automotive market, a short time-to-market cycle means, among other things, that preparation for the production of new models has to be accelerated. Production is located all over the world. For this reason, the

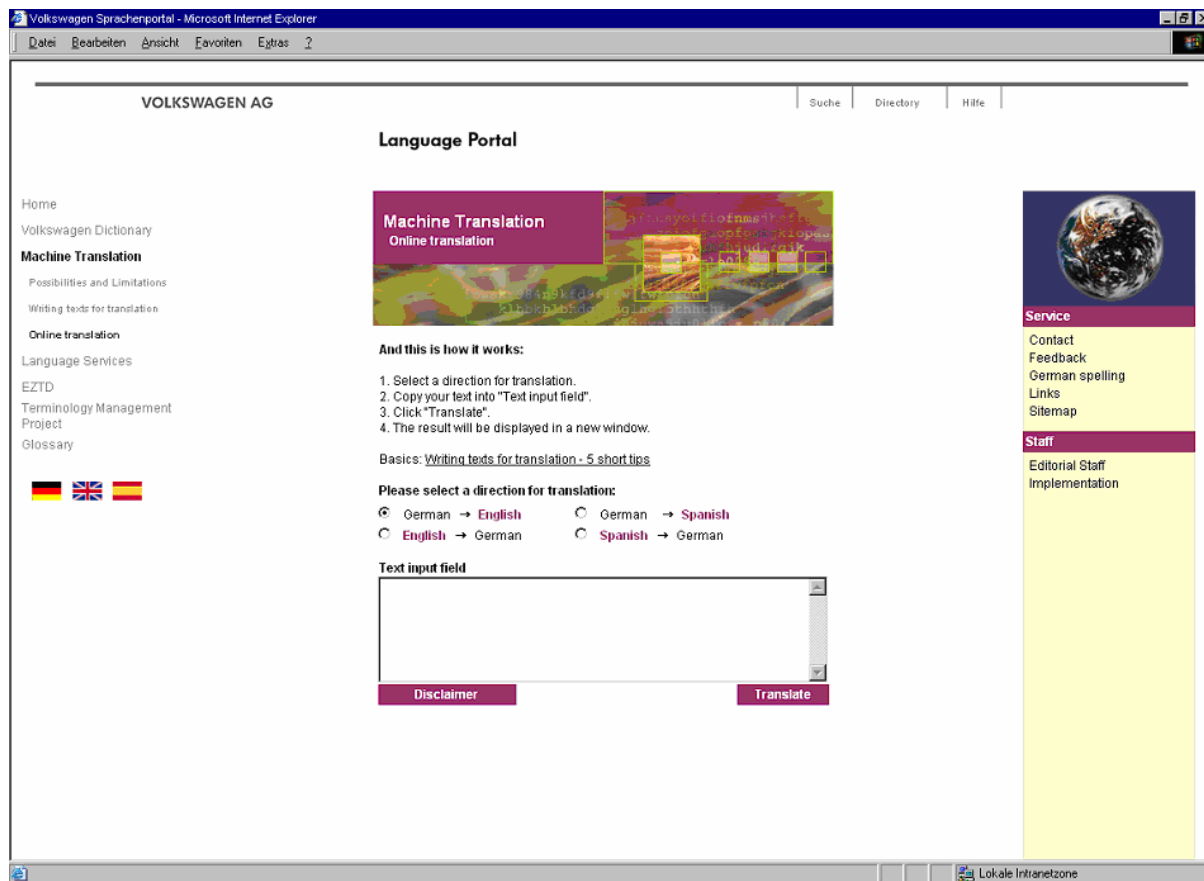


Figure 3. Volkswagen Language Portal – Machine Translation

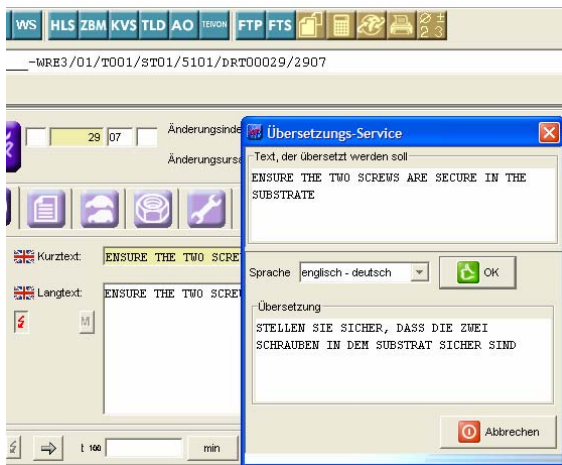


Figure 6. Application “Arbeitsplan”

4. MT Technology for Volkswagen

The most important requirements for the successful realisation of the above projects – apart from language quality – were the availability of an extensible, high-performance translation server, easy integration into existing Volkswagen applications, and guaranteed terminology migration.

4.1. Translation Server

4.1.1. Architecture

The **Compendium Translator Server** is a distributed client-server system that allows communication and sharing of services in a distributed application. Its core, the **Task Scheduler**, allows for configurations involving large numbers of translation engines, and also controls operations in configurations with large numbers of different end-user clients, typically web clients, clients of customer-specific applications or professional translator desktop clients (Distributed Task Scheduler system [Bernardi et al. 2001]). Its main features are:

- Automatic load balancing

- Modularity (core configuration + high level modules)
- Portability (UNIX, Windows NT, Java)
- Service independence (plug & play)
- Security
- State-of-the-art interfacing (see Integration section below)
- Very high performance
- DB independence
- Continual storage and asynchronous processing of large jobs
- Support of user authentication
- Support of reporting, billing
- Flexible setting of translation parameters
- Distributed usage of translation memory
- Starting as service/background processes

All tasks are collected in a **pool** and fed through a special **queue**, allowing for robust and easy task management. The server includes a client for server task administration.

The **pool** accepts tasks from clients, stores them in a memory until requests are processed, and passes results back to the clients. Changes in task states result in events that are passed by the pool (supplier) to a corresponding filtering queue.

A **queue** server implements distinct queues for each state transition. **Translation engines** (or other consumers) register with queues to receive information on changes in the status of tasks in the pool. A pool server now implements three types of task status/events: Waiting, Running, and Done. Components that are interested in specific events register with the corresponding filtering queues. Thus the workload of the pool is further reduced and the overall performance is increased. Depending on the customers needs, the underlying translation engines may contain from one to all **language directions**.

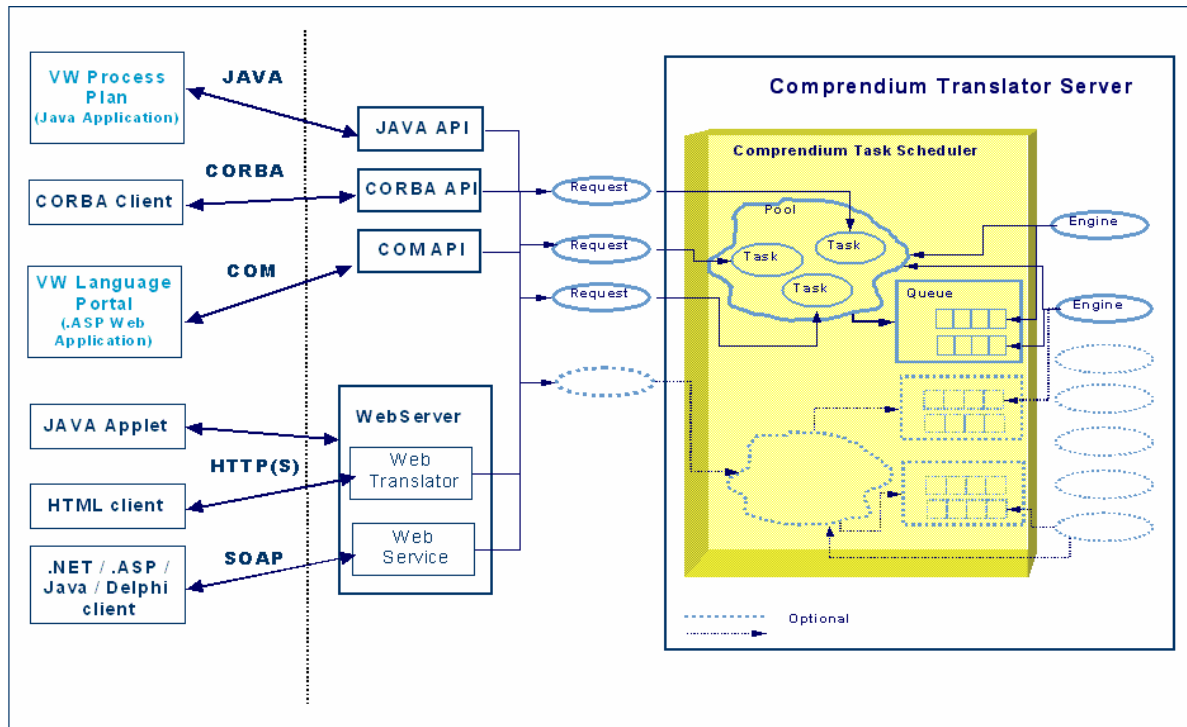


Figure 7. Configuration Overview

The **Compendium Translator Engine** is the translation engine that contains the kernel software, the translation grammars and the lexicons for the language directions included in the system.

In both VW scenarios the Compendium Translator Server is used without any adaptations simply “out of the box”. Three/two MT engines are configured with the server (respectively), see figure 7 (Configuration Overview).

4.1.2. Integration

The Compendium Translator Server is equipped with several APIs (JAVA, CORBA, COM, SOAP, HTTP) that make integration of the translation service in third party applications simple and fast. The Volkswagen Language Portal uses the COM API, the “Process Plan” uses the JAVA API. In both cases the integration required less than one day.

A special adaptation for the “Language Portal” project was required to create a special automatic logging of words unknown to the MT systems sorted by a given VW department. Unknown word logging is an integrated feature of our MT systems which can be switched on or off by a user: Group-specific terminology logging was implemented before installation at VW. This allows control and charge of the ter-

minology coding service by each department involved.

4.2. Terminology Migration

Prior to the realisation of the “Language Portal” and “Process Schedule” projects, Volkswagen and Compendium carried out a pilot project to demonstrate that the migration of specialized terminology significantly improves the translation quality of text in a specific domain.

The most important terms used by Volkswagen in the field of production technology are stored in a terminology database. For the pilot project, Compendium imported approx. 4,000 terms for each translation direction. The database entries were exported to Excel sheets, the corresponding linguistic features (e.g. POS) were replaced by the corresponding Compendium features, and the newly created file (in CSV format, see figure 8) were automatically imported into the MT lexicon by using Compendium’s professional dictionary administration tool LexShop which includes an automatic input parser and defaulter to create the MT system values.

Quality benchmarking showed highly satisfying results, as shown in the following example:

	A	B	C	D	E	F	G	H
1	shock absorber	NST		absorber	Stoßdämpfer	NST		dämpfer
2	coil spring	NST		spring	Schraubenfeder	NST		feder
3	twin tube shock absorber	NST		absorber	Zweirohrdämpfer	NST		dämpfer
4	zinc plate	NST		plate	Zinkscheibe	NST		scheibe
5	base	NST			Unterlage	NST		
6	wishbone	NST			Querlenker	NST		
7	wheel bearing housing joint	NST		joint	Verbindung zum Radlagergehäuse	NST		Verbindung
8	fitting position	NST		position	Einbaulage	NST		lage
9	special tool	NST		tool	Spezialwerkzeug	NST		werkzeug
0	Nm of torque	NST	ABB	Nm	Drehmoment von Nm	NST	ABB	Drehmoment
1	rpm	NST	ABB		1/min	NST	ABB	
2	maximum output	NST		output	maximale Leistung	NST		leistung
3	kW	NST	ABB		kw	NST	ABB	
4	plastic intake manifold	NST		manifold	Kunststoff-Saugrohr	NST		Saugrohr
5	use of plastics	NST		use	Einsatz von Kunststoff	NST		Einsatz
6	manifold	NST			Saugrohr	NST		
7	intake manifold	NST			Saugrohr	NST		
8	throttle valve control unit	NST		unit	Drosselklappensteuereinheit	NST		einheit
9	intake manifold pressure sender	NST		sender	Geber für Saugrohrdruck	NST		Geber
0	intake air temperature sender	NST		sender	Geber für Ansauglufttemperatur	NST		Geber
1	roller-type cam follower	NST		follower	Rollenschlepphebel	NST		hebel

Figure 8. Terminology Import File before defaulting the MT feature values

English Source:

Shock Absorbers The shock absorbers and the coil springs are separate. Twin tube shock absorbers are used. The coil springs have a zinc plate as a base.

Wishbones When the shock absorbers are removed and the vehicle is jacked up, the wishbones sag to such an extent that they can become damaged (cracks) at the wheel bearing housing joint. To prevent this, the wishbone should be fixed in the normal fitting position with a special tool before detaching the shock absorber.

Human German Translation:

Stoßdämpfer Die Stoßdämpfer und die Schraubenfedern sind getrennt angeordnet. Die Stoßdämpfer sind Zweirohrdämpfer. Die Schraubenfedern haben eine Zinkscheibe als Unterlage.

Querlenker Bei demontierten Stoßdämpfern und angehobenem Fahrzeug hängen die Querlenker so stark durch, dass sie an der Verbindung zum Radlagergehäuse beschädigt (Risse) werden können. Um

dies zu verhindern, wird mit einem Spezialwerkzeug der Querlenker in der normalen Einbaulage gehalten, bevor der Stoßdämpfer gelöst wird.

MT German Translation before VW terminology import:

Stoßdämpfer Die Stoßdämpfer und die Schraubenfedern sind gesondert. Zwillings-Röhrenstoßdämpfer werden benutzt. Die Schraubenfedern haben eine Zinkplatte als ein Sockel.

Wishbones Wenn die Stoßdämpfer entfernt werden und das Fahrzeug hochgehoben wird, hängen die wishbones zu solch einem Umfang durch, dass sie (Bruchrisse) an der Radlagergehäusefuge beschädigt werden können. Um dies zu verhindern, sollte das wishbone in der normalen Teilstellung mit einem besonderen Werkzeug befestigt werden, bevor man den Stoßdämpfer abtrennt.

MT German Translation after VW terminology import:

Stoßdämpfer Die Stoßdämpfer und die Schraubenfedern sind gesondert. Zwei-

rohrdämpfer werden benutzt. Die Schraubenfedern haben eine Zinkscheibe als eine Unterlage.

Querlenker *Wenn die Stoßdämpfer demontiert werden und das Fahrzeug angehoben wird, hängen die Querlenker zu solch einem Umfang durch, dass sie (Risse) an der Verbindung zum Radlagergehäuse beschädigt werden können. Um dies zu verhindern, sollte der Querlenker in der normalen Einbaulage mit einem Spezialwerkzeug gehalten werden, bevor man den Stoßdämpfer abtrennt.*

Terminology conversion has now become an ongoing process for MT maintenance management.

5. Outlook

After an encouraging start regarding user acceptance and feedback, the next steps Volkswagen envisages include the introduction of terminology management and authoring tools in other departments and areas that deal with text production, service, customer relationship management, and R&D.

Volkswagen also intends to add more language-relevant applications (i.e. translation) as

well as multilingual content, for example in East European languages and Chinese.

A long-term goal is to turn terminology, and thus language management, at Volkswagen into a Group-wide service that would be an integral part of a multitude of workflows.

6. Acknowledgements

We would like to thank everyone who has contributed to the above projects and to this article, especially Georg Krekeler (Volkswagen, K-PS), Hans-Werner Krause (Volkswagen, K-GOE), Hans-Jürgen Wrensch (Volkswagen, K-GOE) and our former MT colleague Steve McLaughlin.

7. Bibliographical References

BERNARDI, Ulrike, GIESELMANN, Petra, and McLAUGHLIN, Steve (2001). 'A Taste of MALT'. In: 'MT Summit VIII' conference, 18-22 September 2001, Santiago de Compostela, Galicia, Spain.

HERTEL, Helmut; PORSIEL, Jörg (2004): 'Verstehen wir uns?' in: 'DIALOG – Das Service-Magazin von Volkswagen', pp. 120-121, Ausgabe 6, Dezember 2004.

PORSIEL, Jörg (to appear 2005): 'Terminologienagement bei Volkswagen' in: 'Mitteilungen für Dolmetscher und Übersetzer', Heft 3, August 2005.