

Recibido / Received: 25/07/2024  
Aceptado / Accepted: 24/09/2024

Para enlazar con este artículo / To link to this article:  
<http://dx.doi.org/10.6035/MonTI.2025.ne8.19>

Para citar este artículo / To cite this article:

SÁNCHEZ RODAS, Fernando. (2025) "Proper names in German > Spanish tourist translation: the case of Baden-Württemberg." In: Álvarez Jurado, Manuela & Francisco Luque Janodet (eds.) 2025. *La traducción del discurso turístico: desafíos actuales y emergentes / The translation of tourism discourse: current and emerging challenges*. *MonTI* Special Issue 8, pp. 542-568.

## PROPER NAMES IN GERMAN > SPANISH TOURIST TRANSLATION: THE CASE OF BADEN-WÜRTTEMBERG

FERNANDO SÁNCHEZ RODAS

fersanchez@uma.es  
Universidad de Málaga

### Abstract

Works devoted exclusively to proper names in tourist translation are still very scarce. However, both the average tourist and the specialized writer could agree that names are one of the key vehicles of the "double appellation" strategy attributed to tourist discourse by Nobs Federer (1995): inform and excite in order to attract the visitor to a destination. This study compiled a German > Spanish comparable corpus from the official website of the Baden-Württemberg Tourism Department, on which manual Named Entity Recognition (NER) and annotation were performed. Onomastic data were later regrouped automatically using Excel and categorized according to their (non-)mediation schemes: translation, blending, borrowing, and further combinations. Results yielded that LOC (Location) and MISC (Miscellaneous) entities are the most frequent in both languages, but general numbers drop dramatically in Spanish (254 annotated named entities against 901 in the German corpus). The combined use of names with paralinguistic signs such as brackets and inverted commas is also quantified and discussed cross-linguistically.

**Keywords:** Tourist translation; Onomastics; German-Spanish; Toponymy; Baden-Württemberg.



Este trabajo se comparte bajo la licencia de Atribución-NoComercial-CompartirIgual 4.0 Internacional de Creative Commons (CC BY-NC-SA 4.0): <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

## Resumen

La investigación dedicada en exclusiva a los nombres propios en traducción turística es todavía muy escasa. Aun así, tanto el turista medio como el escritor especializado podrían coincidir en que los nombres son uno de los vehículos clave de la doble función persuasiva atribuida al discurso turístico por Nobs Federer (1995): informar y a la vez emocionar para atraer al visitante a un destino. Este artículo describe la creación de un corpus comparable alemán > español del sitio web oficial de la Oficina de Turismo de Baden-Wurtemberg, empleado para anotar entidades nombradas reconocidas de manera manual (NER). Posteriormente, los datos onomásticos se reagrupan automáticamente mediante Excel y se clasifican de acuerdo con sus esquemas de (no) mediación: traducción, fusión, préstamo y otras combinaciones. Los resultados indican que las entidades LOC (topónimos) y MISC (misceláneas) son las más frecuentes en ambos idiomas, pero los números generales disminuyen drásticamente en español (254 entidades nombradas frente a 901 en el corpus alemán). El uso cuantificado de nombres con signos paralingüísticos como paréntesis y comillas también es objeto de debate en ambas lenguas.

**Palabras clave:** Traducción turística; Onomástica; Alemán-español; Toponimia; Baden-Wurtemberg.

## 1. Introduction

The influence of John Stuart Mill's philosophy of language from 19th century to the present day is largely responsible for the fact that names are, even today, an anecdotal object of study in Linguistics and a marginal one in Translation and Interpreting. Millian theory strips proper names of any kind of connotation and considers them non-significant marks, similar to library bookplates, whose only purpose is to identify the books to which they are attached (Mill 1875: 37). This utilitarian linguistic stance was present in translation studies during most of the 20th century, the most radical voices claiming that the main difference between common and proper nouns lay in the practical untranslatability of the latter (Mańczak 1969: 290). From the 1980s onwards, discordant voices began to emerge, advocating the possibility of translating proper names in a number of discourses: literary (Hermans 1988), legal and financial (del Águila 1997), journalistic (Moya 2000), and audiovisual (Cuéllar Lázaro 2016).

Tourist translation potentially shares traits with some of the above fields, but works devoted exclusively to proper names are still very scarce. Both the average tourist and the specialised writer could agree that names are one of the key vehicles of the “double appellation” strategy attributed to tourist discourse by Nobs Federer (1995): inform and excite in order to attract the visitor to a destination. However, this initial assumption should be confirmed with empirical data, due to the relevance of names for the translator mediating between tourist texts. To pursue such end, this article presents a state of the art on name translation in tourist texts (Section 2) before elaborating a brief theory on the topic from an interventionist, constructional, and semiotic perspective (Section 3). Section 4 describes a named entity recognition method for the empirical study of name translation in a corpus of German > Spanish tourism websites. Section 5 comments on the extracted data, trying to determine which name categories are translated or not in the chosen corpus, in what form and how often. The conclusions in Section 6 include a general comparison with the results from previous studies on institutional name translation.

## 2. State of art

Few publications in tourist translation involving Spanish explicitly use terms such as “proper name”, “toponym” or “anthroponym”, understood as linguistic constructions fully integrated into the grammatical, semantic and discursive layers of the source and/or target languages (Sánchez Rodas 2024). References to the topic are usually integrated within broader objectives.

In the Spanish > English translation of flamenco tourism texts, Venceslá Mariscal (2005) detects problems with anthroponyms and toponyms that can be extrapolated to other types of tourist texts, as well as to general and specialized translation fields. For Durán Muñoz (2011), proper names are one of the greatest challenges faced by tourism translators in their mediator role. Names abound in tourism texts because of their frequent references to historical facts and figures, as well as monuments and institutions of the source culture (Durán Muñoz 2012a; 2012b). Fuentes Luque (2016) mentions proper names as important marketing elements in the translation

of tourism advertising. Nevertheless, they are difficult units for the human translator, and machine translation is still not able to provide enough help with them (Fuentes Luque & Santamaría Urbieto 2020). Their difficulty could also be a reason why source names are sometimes omitted. In a series of Spanish > Russian travel guides, Polyakova (2021) perceives an absence of relevant place and monument names in the Russian translations.

It is true that several translation studies (including tourism translation) cover names under the umbrella of “culturemes”, “cultural terms” and similar denominations (Ramírez Almansa 2019). They are not addressed in this article, because such classification can lead to confusion for a number of reasons. Devising an artificial category that separates what is culture in language from what is not dissociates human communication, even more in the case of proper names, since all languages are cultural vehicles and expressions in their own. The word “culture” and its derivatives represent extremely generic concepts prone to serve as a catch-all for the scholar and could therefore lead to erroneous generalizations (Minkov 2013). For example, it is not a rule of thumb that place names, names of dishes, dances and typical events of a culture should not be translated in tourism, and that a descriptive translation should always be provided (Soto Almela 2013). Rather, this would depend on a combination of numerous factors including language pairs and direction, production mode (oral or written), mediation mode (translation, interpretation, and their variants) and mediation context. The casuistic of name translation can go as far as to require different solutions depending on the individual features of each name, such as its institutionalization degree (Allerton 1987).

### 3. Proper names in tourist translation

#### 3.1. *Toponyms*

In tourist translation, names require a particularly demanding mediation exercise, since translators have to intercede in two ways. They must ensure that the target community has access to knowledge which may be specific even for the native reader, and such knowledge must be also presented in an attractive or suggestive form. It is easy for the translator of tourism to

feel a tension in names between the desire to transmit knowledge and the need of the client to produce an economic return, and this must be resolved through concrete choices. The specificities of each name require individual negotiations, which explains the multitude of techniques and forms used in tourist texts: names are sometimes translated, sometimes borrowed, or translated and also described between brackets. The solution, as seen in Polyakova (2021), can even consist of their removal. The following sections synthesize some constructions that translation scholars propose for names in tourist texts, together with examples. They can be subdivided into the onomastic triad of toponyms, anthroponyms and chrematonyms, which comprises most proper names as traditionally considered.

Newmark (1988) mentions toponyms as characteristic elements in tourist guide translation. They can be transferred through a combination of translation and borrowing. Kelly (2005: 166) documents deeper transformations (*Costa Tropical* → *Eastern Costa del Sol*) and suggests reflecting on their extra-linguistic consequences<sup>1</sup>. Durán Muñoz (2012a: 110, translated) studies the translation of toponyms combined with geographical nouns such as lakes, mountains, hills, etc., with up to three possibilities:

1. Names that have a translation equivalent in the target language because of their frequency and popularity in both cultures, such as *rio Támesis* for the Thames River.
2. Names where geographical features are an integral part, such as *Sierra Nevada*. Here, the name should be kept and an equivalent of the feature be added for clarification: ‘mountain range Sierra Nevada.’
3. Names which also integrate a geographical feature but it is translatable through a target language equivalent, such as *lago Lomond* for Loch Lomond.

Another construction similar to the above is found by Venceslá Mariscal (2005: 148-149) in Spanish tourist guides on flamenco. The *palos* (different musical traditions within flamenco) are expressed in original Spanish with a noun phrase composed of at least two elements: the term designating the *palo* and their placename of origin (*alegría de Córdoba*, *taranta de Linares*, *fandango de Lucena*). This construction is replicated in several Spanish

---

1. A similar example in German > Spanish could be translating *Baden-Württemberg* as *la región de la Selva Negra*, ignoring its further geographical landmarks.

tourist texts on gastronomy (*vino de Montilla*, *cerezas del Jerte*), architecture (*castillo de Almodóvar*, *palacio de Aranjuez*), and other topics. In the context of German < > Spanish translation, German employs similar syntactic groups but these do not work exactly the same. Instead of prepositions plus toponyms, German usually employs an adjective in capital letters derived from a placename, and this adjective becomes an integral part of a wider proper name (*Heidelberger Schloss*, *Bayerischer Wald*)<sup>2</sup>. At any case, a brief description of the location may be added if the toponym is little known among the target Spanish-speaking community.

### 3.2. *Anthroponyms*

For Durán Muñoz (2012a), the most frequent anthroponyms in tourist translation are historical and artistic figures, like painters, writers, musicians, kings and popes. As with toponyms, the German > Spanish direction poses challenges in this category. Long-standing, coined translations (*Ludwig I von Bayern* > *Luis I de Baviera*, *Albrecht Dürer* > *Alberto Durero*) coexist with the current borrowing tendency (*Angela Merkel*, *Sebastian Kurz*).

### 3.3. *Chrematonyms*

Chrematonyms include proper names of social phenomena such as holidays or historical events (*Tag der Deutschen Einheit*), public or private organizations (*Bundesministerium für Gesundheit*, *Siemens*) and individually or mass-produced objects (*Titanic*, *Haribo*, *Also sprach Zarathustra*). A set of chrematonyms is a “chrematonymy”, and their field of study is called “chrematonomastics” (Gałkowski 2017). Despite their social relevance, chrematonyms are rarely addressed in tourist translation studies, probably because of their marginal status in Western Onomastics (but not in Central and Eastern Europe academia).

---

2. German adjectives in *-er* derived from toponyms usually refer to towns or cities (Heidelberg in this case). In fact, they are always written in capital letters, even if they are not part of a proper name (e.g. *ich liebe Heidelberger Wein!*). See Section 4.2. below for more details.

Durán Muñoz (2012a) refers to the frequent occurrence of names of institutions and paintings in the tourist source text, which are generally translated (*Adam und Eva* > *Adán y Eva*). According to Särkkä (2007), museum names are often at the frontier between appellation and onymy. They are usually constructed without the help of any prototypical proper name and translated the same way (e.g. the Viennese *Kunsthistorisches Museum* > *Museo de Historia del Arte*), although there are exceptions such as museum houses (*Friedrich-Ebert Haus*)<sup>3</sup>. In museums and cultural centers, the visitor is sure to find numerous references to historical events and dates, for which Newmark (1988) and Moya (2000) propose a wide range of possibilities ranging from borrowing to descriptive translation. The solution will depend on whether a coined translation exists (*Wiener Kongress* > *Congreso de Viena*), or on the target text specialization (*Anschluss* can remain the same for Spanish-speaking readers educated on the topic, or else be replaced by a description such as *la anexión nazi de Austria*).

## 4. Methodology

### 4.1. BADENCOR corpus

The methodology of the present study is divided into three parts. It proposes an empirical method in chronological steps for the bilingual study of names with pilot or ad hoc corpora. The method includes the compilation of a small corpus of tourist websites (4.1.), a fully manual recognition of the named entities they contain (4.2.), and the automatic grouping of the data via Excel before their analysis (4.3.).

---

3. *Kunsthistorisches Museum* is translated in private guiding companies such as Civitatis (<https://www.disfrutavienna.com/museo-historia-arte>). However, it is borrowed and translated in the Vienna Tourist Office (<https://www.wien.info/es/artey-cultura/museos-y-exposiciones-en-viena/museos-mas-destacados/kunsthistorisches-museum-wien-museo-de-historia-del-arte--350442>), and just borrowed in the official website translation into Spanish (<https://www.khm.at/informacion-para-visitantes/>). This is another example of different solutions to names in accordance to target text specialisation. Specialisation upgrades usually involve a progressive shift from full translation to borrowing.

In order to observe tourism names in non-translated German and translated Spanish, an ad hoc corpus called BADENCOR was compiled. The corpus contained promotional web pages from the Tourism Department of the German region of Baden-Württemberg (2024). Of all the *Länder* in Germany, Baden-Württemberg is presented as a convenient choice for studying tourist discourse in the German > Spanish language pair, due to the potential for Spanish stays in the region. In 2019, the Baden-Württemberg Tourist Board, together with the Tourist Offices of the Upper Black Forest and Baden-Baden, organized a joint promotional presentation in Madrid with the support of the German National Tourist Office and the German Embassy in Spain. At that meeting, the officers emphasized that the destination recorded 340,000 overnight stays by Spanish travelers in 2018 and they were confident that this figure would increase in the coming years (Marina 2019). Although these figures are pre-pandemic, Baden-Württemberg still is attractive enough to seduce Spanish tourists with a medium or medium-high purchasing power: quality food and wine, thermal and environmental tourism, and historical towns with recognized heritage such as Baden-Baden, Freiburg, or Heidelberg. Allegedly, Baden-Württemberg is also the German region with the most hours of sunshine, which led public tourism promoters to adapt the original slogan *Wir Sind Süden* to the Spanish *Todos los colores de Alemania*<sup>4</sup>.

BADENCOR is a bilingual comparable German-Spanish corpus of 18 texts (9 per language) written to highlight the tourist attractions of some of Baden-Württemberg's most appealing cities: Baden-Baden, Freiburg, Heidelberg, Heilbronn, Karlsruhe, Mannheim, Pforzheim, Stuttgart and Ulm. These texts have their own section on the website under the headings *Städte des Südens/Ciudades*. Text about cities were chosen over text about regions and natural attractions, because they were expected to be

---

4. In English and French, similar strategies are perceived with the slogans "The Sunny Side of Germany" (<https://www.tourism-bw.com/>) and *La chaleur de l'Allemagne du Sud* (<https://www.tourisme-bw.fr/>). The adaptation is linguistic and graphic, as each language slogan has its own logo.



more balanced in terms of toponymy, anthroponymy, and chrematonymy distribution<sup>5</sup>.

Although the number of texts in both languages is even, the corpus is comparable instead of parallel. The Spanish texts are considerably shorter than the German versions and organize the information differently. Changes are also noticeable in the user's interaction with the website. The German version includes a sliding menu of tourist attractions with individualized hyperlinks and a preview of related articles at the bottom of each text, but this is absent from the Spanish versions. Thus, Baden-Württemberg illustrates a comprehensive tourism mediation strategy that not only includes language transfer, but also a selective screening of the original information in German. The purpose of this curation is to maintain only the most interesting or well-known references for the Spanish visitors (Figures 1 and 2).

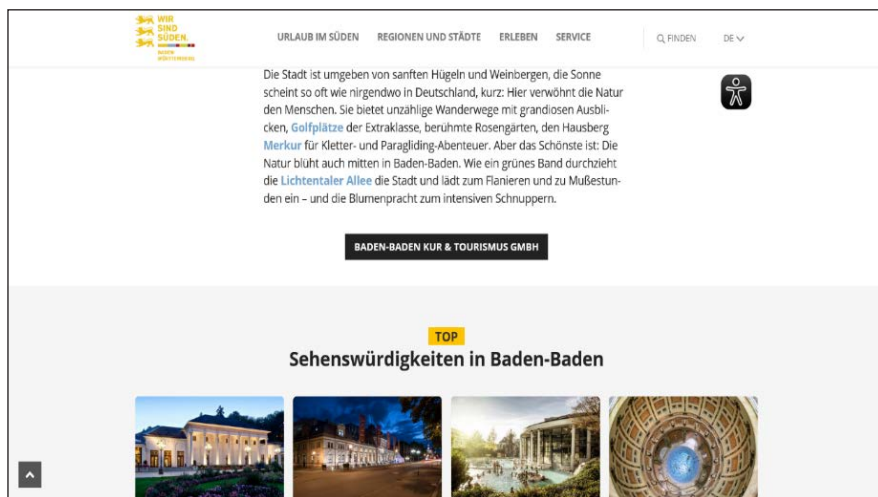


Figure 1. Official Baden-Baden tourism website in German with sliding menu below

5. It is still possible to expand the corpus in the future with regional sections like the Black Forest, and/or with similar texts in tourist websites from other *Länder*.

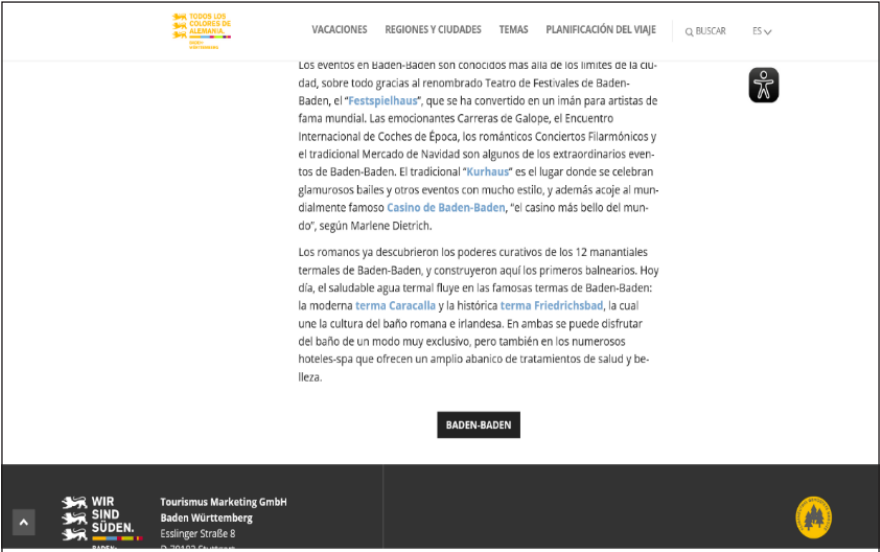


Figure 2. Official Baden-Baden tourism website in Spanish<sup>6</sup>

The disparate length of the texts results in size differences between the subcorpora, calculated using Sketch Engine (Table 1). Sizes are based on TXT UFT-8 versions of the web pages without hyperlinks, alternate text, and other concealed HTML content. These were deleted to make analyzed texts as faithful as possible to what average visitors can directly read on the website.

6. The sliding menu is absent, and the text body is rearranged. The top-left logo is changed to Spanish (*Todos los colores de Alemania*) but kept in German at the bottom left (*Wir sind Süden*).

BADENCOR CORPUS				
Subcorpus	Documents	Sentences	Tokens	Unique words
N-T_DE (Non-Translated German)	9	293	7,905	2,367
T_ES (Translated Spanish)	9	67	2,079	716
	18	360	9,984	3,083

Table 1. Size of the BADENCOR corpus in documents, sentences, tokens and unique words (Sketch Engine calculations)

The original corpus is up to 73.7% larger in tokens than the translated one (5,826 tokens of difference). There is also a considerable down-size in the number of unique words in Spanish (716, 69.7% less than in German). These substantial changes may be correlated with the cross-lingual adaptation strategies of this website, the German > Spanish pair, tourist translation in general, or a combination of all. The intermodal corpus PETIMOD 1.0., previously compiled for the study of English > Spanish name translation and interpreting in institutional texts, had similar dimensions but a more balanced number of tokens (Corpas Pastor & Sánchez Rodas 2022).

4.2. Manual named entity recognition (NER)

Once the bilingual corpus was compiled, the next methodological step was the annotation of proper names in both languages. Annotation was carried out according to the hypothesis that proper names behave differently from common lexicon and terminology and require their own identification and computation methods. Natural Language Processing (NLP) has contributed in recent years to the development of a suitable tool for this endeavor, which is Named Entity Recognition (NER).

NER can be conceived as a utility or a task. Every NER starts from a corpus of texts and aims at identifying and categorizing Named Entities

(NEs). NEs are linguistic units which designate key information and unique objects, be them real or fictitious. From the point of view of the end-user, NER is usually an automatic process performed through a dedicated software or interface. Nevertheless, human researchers should also carry out manual NER in order to improve automatic results and amend possible errors, such as homographs, orthographic variants, morphological inflection and multiword entities (Jacquet *et al.* 2019). Manual NER consists of reading the texts in search of possible NEs and annotating them according to a set of labelling rules, the spelling rules of the language and contextual information which suggests a very specific person, place, organization, etc. This task can lean on resources such as spelling handbooks, monolingual encyclopedic dictionaries, onomastic dictionaries, or lexical dictionaries indexing proper names.

In the present study, a manual-only NER was performed for two reasons: the small size of BADENCOR corpus and the didactic value of manual recognition for tourist genres, where proper names are still not explored in depth. The annotation has been carried out directly in the corpus' TXT files, square-bracketing the entities detected in German or Spanish and including one of four tags: PER: (names of people), ORG: (names of public and private organizations), LOC: (geopolitical and geographical locations) and MISC: (miscellaneous entities, such as names of documents, festivities or works of art). Although there are models with more categories, PER, ORG, LOC and MISC are among the most basic tags for corpus-based entity annotation. Table 2 below provides more details on the entities covered by each category, with examples extracted from the corpus<sup>7</sup>.

---

7. Ambiguous names that could refer to both a building and its hosted institution (e.g. museums or restaurants) have been labelled LOC because the main function of the text is that the tourists visit those places. This technique also avoids mistakes in case the institution no longer exists or has an official name different from the building, which is harder to document. ORG have been reserved to cases where visitation is not that common, or to official denominations abstractly comprising several buildings (*Siemens, Bundesverfassungsgericht, Universidad Ruperto Carola*).

LABEL	DESCRIPTION	CORPUS EXAMPLES
PER:	Real or fictitious people	<i>Johannis Reuchlin, Picasso</i>
ORG:	Public and private organisations (companies, agencies, institutions, etc.)	<i>Siemens, Universidad Ruperto Carola</i>
LOC:	<ul style="list-style-type: none"><li>– Non-geopolitical locations (mountains, rivers, etc.)</li><li>– Geopolitical locations (cities, regions, countries, etc.)</li><li>– Buildings, airports and bridges</li></ul>	<i>Neckaruferpark, Europa, Villa del Vino</i>
MISC:	<ul style="list-style-type: none"><li>– Nationalities, religious or political groups</li><li>– Objects, vehicles, food and other products</li><li>– Hurricanes and other named social phenomena (battles, wars, sporting events etc.)</li><li>– Titles of books, songs and paintings</li><li>– Named documents turned into laws</li><li>– Language names</li><li>– Dates or periods, absolute or relative</li><li>– Time indications shorter than one day</li><li>– Percentages (including the % symbol)</li><li>– Monetary values and units</li><li>– Measurement values such as weight or distance</li><li>– Ordinal numbers</li><li>– Generic cardinal numbers</li><li>– Other named entities</li></ul>	<i>Bundesgartenschau 1975, Hans-im-Glück Brunnen, Freiburger, 800 hectáreas, Carreras de Galope</i>

Table 2. Description and corpus examples of manual NER categories

For Spanish native speakers, the manual recognition of named entities in German can be a task of great difficulty. While Spanish normally uses an initial capital letter to differentiate proper names from common nouns, the spelling rules of German require initial capital letters in both cases. From the Spanish speaker’s point of view, this equalizes both forms and may lead to wrongly assign onymic values to an appellative form, or vice versa. However, German spelling and grammar books offer their own explanations for the recognition and formalization of proper names, which comes handy for manual NER. The Basic Duden Handbook (2014: 45) prescribes

the initial capitalization of proper names (*Eigennamen*). This rule applies to anthroponyms (*Heinrich Heine*), toponyms (*Europa*, *Mannheimer Straße*), constellations and celestial bodies (*Kleiner Bär*, *Mars*), public and private institutions (*Vereinte Nationen*), historical events (*der Westfälische Friede*) and names of offices and titles (*der Heilige Vater*, *der Regierende Bürgermeister*), among others. For anthroponyms and toponyms, the handbook details specific spelling cases (Table 3):

ANTHROPONYMS	
Some word categories that are part of the anthroponym, such as adjectives, participles, pronouns and ordinals, are capitalized.	<i>Der Alte Fritz</i> , <i>Katharina die Große</i> , <i>Albrecht der Entartete</i> <i>Unsere Liebe Frau</i> ( <i>Maria als Mutter Gottes</i> ) <i>Elisabeth die Zweite</i> , <i>Karl der Achte</i> , <i>Ludwig der Erste</i>
Adjectives derived from anthroponyms are usually written with an initial lower case. However, it is also possible to write them with a capital letter and use an apostrophe (') to separate them from the adjective ending, if the aim is to highlight the name of the person.	<i>einsteinsche Relativitätstheorie</i> ( <i>Einstein'sche Relativitätstheorie</i> ) <i>freudsche Fehlleistung</i> ( <i>Freud'sche Fehlleistung</i> ) <i>grimmische Märchen</i> ( <i>Grimm'sche Märchen</i> ) <i>luthersche Bibelübersetzung</i> ( <i>Luther'sche Bibelübersetzung</i> ) <i>ohmsches Gesetz</i> ( <i>Ohm'sches Gesetz</i> )
TOPONYMS	
Some word categories that are part of the toponym, such as adjectives and participles, are capitalized.	<i>Der Bayerische Wald</i> , <i>das Rote Meer</i> , <i>der Indische Ozean</i> , <i>der Stille Ozean</i> , <i>das Kap der Guten Hoffnung</i> , <i>die Vereinigten Staaten von Amerika</i>
All adjectives ending in <i>-er</i> that derive from toponyms are written with an initial capital letter.	<i>Der Hamburger Hafen</i> , <i>der Schweizer Käse</i> , <i>das Wiener Schnitzel</i>
All adjectives ending in <i>-isch</i> that are derived from toponyms are written with a lower case, provided they are not part of a proper noun.	<i>der badische Wein</i> , <i>der holländische Käse</i> , <i>der indische Tee</i> , <i>das russische Roulette</i> (but: <i>die Mecklenburgische Seeplatte</i> )

Table 3. Initial capital and lowercase letters of German anthroponyms and toponyms (adapted from Duden 2014: 46)

If we apply Table 3 to NER, there are some German units (mostly adjectives and participles) capitalized as part of multi-word NEs, such as *Katharina die Große* or *Mecklenburgische Seeplatte*. When it comes to adjectives ending in *-er*, some can be considered part of a multiword LOC (*Hamburger Hafen*) and others are single-word MISC (*Schweizer* in *Schweizer Käse*). Some gentils in lower case, such as *badische* in *der badische Wein*, could also be recognised as single-word MISC. Figure 3 shows one NER manual annotation from the BADENCOR corpus based on the Duden spelling recommendations.

```
Menü
    * Urlaub im Süden
    * Regionen und Städte
    * Erleben
    * Service

finden

[MISC:de]

Die kleinste Weltstadt
[LOC:Baden-Baden]

    * Home
    * Regionen und Städte
    * Städte des Südens
    * [LOC:Baden-Baden]

Selten ist es so einfach, eine Stadt im Herzen der Natur zu finden, die
sowohl Entschleunigung als auch Unterhaltung bietet. Doch genau das ist
[LOC:Baden-Baden] – die internationale Bäder- und Kulturmetropole ist
wunderschön idyllisch am Fuße des [LOC:Schwarzwaldes] gelegen und begeistert
Besucher aus aller Welt mit ihrem [MISC:mediterranen] Flair, das auf das
kulturelle Angebot einer Großstadt trifft. Hier fühlen sich Kunstfreunde
und Gourmets genauso wohl wie Naturbegeisterte und Wellnessliebhaber.

stadtansicht von [LOC:Baden-Baden] | © [ORG:Baden-Baden Kur & Tourismus GmbH]
```

Figure 3. Fragment of a manual NER performed on a TXT file

#### 4.3. Data Grouping and Transformation in Excel 365

After manually annotating the named entities in German and Spanish, onomastic data needed to be extracted and grouped to achieve a clear

visualization for analysis. Since the annotated files were plain text, there was a high versatility to choose the data analysis tool. Microsoft Office 365 was selected, more specifically its data retrieval and transformation functions offered by the latest version of Excel 365 (Tab “Data” > “Retrieve and transform data”).

Excel 365 allows the user to embed textual or numerical data from one TXT file in the form of a table. It also allows joint data extraction from all files in the same folder. When choosing this option, Excel 365 automatically adds the original file of each value as metadata. This is especially interesting for corpus work, as it allows to analyze the textual distribution of the values at a glance. The grouping steps for annotated onomastic data were done once per language (Figure 4):

1. Mass data collection from the folders where the annotated sub-corpora were hosted.
2. Filtering of blank rows.
3. Extraction of rows with text between the delimiters [] to retrieve the annotated NEs
4. Splitting of columns by the delimiter : to separate the labels PER: ORG: LOC: and MISC: from the actual named entities.

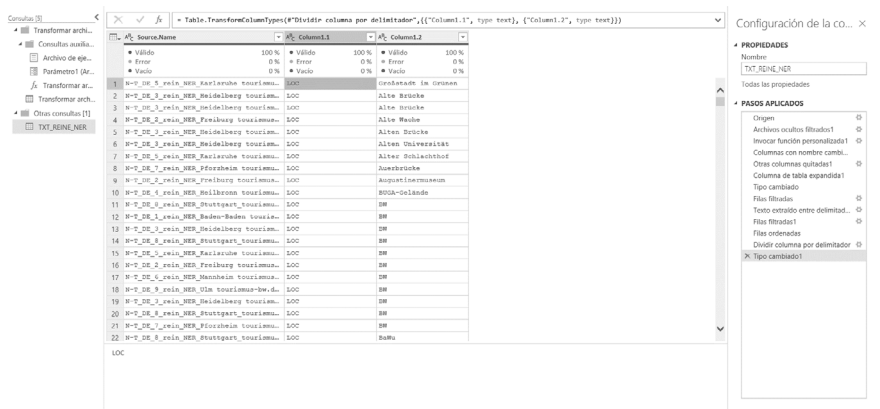


Figure 4. Data transformation of German annotated NEs in Excel 365. Steps are listed on the right side



These steps produced two Excel tables (one per language) with three columns each: NE, NE category and source text. To proceed with the analysis, it was only needed to add further columns specifying the type of mediation (translation, borrowing, etc.).

## 5. Results

### 5.1. *Non-translated German*

In our previous research on organization names in English and Spanish institutional discourse, it was detected that, both in original and translated texts, these units responded to pragmatically motivated schemes or constructions, which could be mediated or non-mediated (Sánchez Rodas 2023). These categories were proposed on the basis of mainstream translation and cognition concepts, and they can be extrapolated to tourist discourse and non-institutional proper names. Non-mediated names are expressed by native speakers in their mother tongue and are an established part of its cultural and symbolic heritage (*Sevilla*, *Museo del Prado*, *Rosalía de Castro*). In contrast, mediated names are alien to this cultural heritage and a result of a certain degree of intervention (mediation) by a third party. Such mediation may adopt different forms: translations (*München*), borrowings (*Oktoberfest*), blends of source and target elements (*República de Weimar*), and more complex combinations. Certain punctuation marks such as parentheses can be considered part of written constructions, especially in mediated names like “*Festspielhaus*” (*Teatro de Festivales*). The analysis of German > Spanish tourism names follows a similar line, uncovering different quantitative distributions of constructions previously seen in institutional discourse as well as describing new constructions marked by the employment of punctuation signs.

In the non-translated German corpus 901 named entities have been labelled (Table 4). The most numerous category is LOC (557), followed by MISC (241), PER (54), and ORG (49) at a greater distance. The prevalence of LOC entities in tourist discourse is almost self-explanatory. In

the BADENCOR corpus, they include names as varied as regions adjacent to Baden-Württemberg (*Bayern*), wine-growing areas (*Baden-Badener Rebland*), mountains (*Königstuhl*), villages (*Bad Herrenalb*), rivers (*Enz*), restaurants (*Alte Wache - Haus der badischen Weine*), museums (*Haus der Stadtgeschichte*), historical buildings (*Historisches Kaufhaus*), and so on. The equally high number of MISC entities comes from a combination of two factors. On the one hand, there are many functional numbers for tourists, such as opening and closing times (09:00 - 17:00), telephone numbers (+49 711 238580, the telephone of Baden- Württemberg' tourist office), height of mountains and width of land areas (600 Meter, 800ha) or inhabitants in a town (125,000). On the other hand, other MISC entities are more oriented to engage visitors, such as references to past ages (18. Jahrhundert), historical events (2. Weltkrieg), annual local festivals (Heidelberger Schloss-Illumination), monuments (Großherzog Karl Friedrich Denkmal), old-fashioned means of transport (Turmbergbahn) and exhibitions or objects unique to a city (Astronomische Kunstuhr). In the PER and ORG categories, almost even in number (54 and 49), there are names of historical figures (Kleist, Lew Tolstoy), as well as local tourist offices (Ulm/ Neu-Ulm Touristik), music orchestras (Württembergisches Kammerorchester Heilbronn) and educational and legal institutions (Ruperto Carola, Bundesverfassungsgericht).

	PER	ORG	LOC	MISC	TOTAL (TYPES)
MEDIATED CONSTRUCTIONS	6	26	14	11	57
BLENDING	0	25	4	4	33
BORROWING	3	0	6	4	13
„BORROWING“	0	0	1	1	2
“BORROWING”	0	1	0	0	1
„BORROWING – TRANSLATION“	0	0	0	1	1
“BORROWING – TRANSLATION”	0	0	0	1	1
TRANSLATION	3	0	3	0	6
NON-MEDIATED CONSTRUCTIONS	48	23	543	230	844
NON-TRANSLATION	48	21	536	226	831
„NON-TRANSLATION“	0	2	6	1	9
“NON-TRANSLATION”	0	0	1	0	1
»NON-TRANSLATION«	0	0	0	1	1
„NON-TRANSLATION“	0	0	0	1	1
‘NON-TRANSLATION’	0	0	0	1	1
	54	49	557	241	TOTAL: 901

Table 4. Mediated and non-mediated NE construction in BADENCOR (non-translated German)

In mediated constructions (53/901), there is a strong influence of English on the German tourism discourse. The most numerous constructions are blending (33). Most blends fusion German and English (*Ulmer City*, *Stuttgart-Marketing*), although there is a single case of a science museum

with a Latin name (*Experimenta Heilbronn*)<sup>8</sup>. Among the borrowings (15/53) English continues to appear, but not hegemonically (*Science Dome*, *UNESCO City of Music*). There are three French borrowings as a consequence of Baden-Württemberg's proximity to France (the cities of *Colmar* and *Mulhouse* and the *Avantgarde* movement), as well as one Spanish name (*Venezuela*). Translations (6/53) are the scarcest mediation type and include naturalized and transliterated anthroponyms and toponyms (*Straßburg*, *Fjodor Dostojewski*).

As in non-mediated constructions, quotation marks play a relevant role in a subgroup of borrowings, where pragmatic signaling (the food festival „*Plaza Culinaria*“) seems to precede spelling requirements („*Lonely Planet*“ as the well-known title of a travel guide). A more complex example can also be observed in the name of an aquatic exhibition, where an English borrowing is hyphenated with a free translation into German: „*Great Barrier Reef - Wunderwelt Korallenriff*“. A very similar instance differs only in the use of capital letters and inverted commas: „*GREAT BARRIER REEF – Wunderwelt Korallenriff*“. This example reinforces the hypothesis of spelling and orthotypographic variations at the service of individualized pragmatic purposes within names.

## 5.2. Translated Spanish

The translated Spanish corpus contains 254 annotated NEs and meets expectations in the distribution of mediated and non-mediated constructions (Table 5). The distribution is reverse to the original texts, and mediation (232) becomes more numerous than non-mediation (22). The proportion of categories remains the same, with LOC (163) once again being more abundant, followed by MISC (59), PER (17) and ORG (15). The greater differences between the two corpora are found in the drastic reduction of the NE number in Spanish (647 fewer than in German) and the increased complexity of the cross-linguistic mediation. Spanish texts have 232 mediations divided into 10 different schemes, compared to 7 schemes in German.

---

8. *experimenta* is written in lower case in the German texts despite being a proper noun. This is probably a branding strategy and was not originally thought to emphasise the Latin origin of the word.

Borrowing (91/232) is more frequent than full translation (78/232), and full translation more frequent than mergers (45/232).

	PER	ORG	LOC	MISC	TOTAL (TYPES)
MEDIATED CONSTRUCTIONS	16	15	162	39	232
BLENDING	0	1	41	3	45
“BLENDING”	0	0	1	0	1
BORROWING	16	10	65	0	91
“BORROWING”	0	1	3	1	5
“BORROWING”+ TRANSLATION	0	0	1	0	1
BORROWING, TRANSLATION	0	0	1	0	1
EXPLICITATION + BORROWING	0	0	2	0	2
TRANSLATION	0	3	43	32	78
TRANSLATION + “BORROWING”	0	0	3	1	4
TRANSLATION + (“BORROWING”)	0	0	0	1	1
NON-MEDIATED CONSTRUCTIONS	1	0	1	20	22
NON-TRANSLATION	1	0	1	20	22
	17	15	163	59	TOTAL: 254

Table 5. Mediated and non-mediated NE construction in BADENCOR (translated Spanish)

The preponderance of borrowing is in line with the diachronic tendencies of Franco Aixelá (2000) regarding English > Spanish names. Translations such as *Heidelburgo* and *Carlos Guillermo de Baden-Durlach* would have been acceptable some time ago, but *Heidelberg* and *Karl-Wilhelm von Baden* remain in German today. In the ubiquitous letterhead of the Baden-Württemberg

tourist office, the official name of the company (*Tourismus Marketing GmbH*), the address (*Esslinger Straße 8*) and the name of the region (*Baden Württemberg*) are also retained in German<sup>9</sup>. Other borrowings are more dubious, as they would be easier to naturalize (*Neckar* instead of *Néckar* or *Napoleon* instead of *Napoleón*, which is a spelling mistake). Translations and blends still concern some toponyms, but these are generally smaller entities, such as neighborhoods (*Barrio de los Pescadores*, *el nuevo barrio Neckarbogen*), museums (*Galería Nacional*, *Museo Engelhorn*) and squares (*Plaza de la Catedral*, *Plaza del Palacio de Stuttgart*), among others.

The rest of Spanish cases represent a wide dispersion of new procedures, combined procedures, and variations through inverted commas. These include “BORROWING” (*“Bauhaus”*), TRANSLATION + “BORROWING” (*iglesia-palacio San Miguel “Schlosskirche St. Michael”*) and EXPLICITATION + BORROWING (*parque Luisenpark*). The last two cases are distinguished from blending by their reiterative intention, which is accentuated in explicitation: the translator extracts a lexical element from the source name and repeats it in Spanish, as if to reinforce the classification of the place (*Plaza Eugensplatz*). As in the original German corpus, the use of inverted commas is inconsistent and could signal different intensities in the mediation process (pragmatic use). Besides, all the inverted commas are English instead of Latin, ignoring prescriptive Spanish spelling.

## 6. Conclusions

This paper sought to expand the rare scientific production on proper names in translation and interpreting studies, using a non-English-centric language pair with a medium level of resources (German > Spanish) and a non-prototypical field (tourist translation). It followed a corpus-based methodology combined with manual named recognition on a small bilingual compilation of official tourism websites. This method yielded a total of 1,155 named entities, with 19 different types of constructions. The added results from EU texts and tourist texts show a cross-gender and

---

9. *Baden Württemberg* (9 cases) loses the German hyphenation and coexists with the isolated variations *Baden-Wurtemberg* and *Wurtemberg* (one case each). They all diverge from the ideal naturalization *Baden Württemberg*.

cross-lingual complicity of proper names with non-linguistic (or not fully linguistic) modes of communication, such as paralinguistic signs. This relation was also explored in recent research on short-form audiovisual translation, where names are an important layer of multimodal communication in social network videos, and even co-built with auxiliary codes such as emojis (Sánchez Rodas & Del Río López 2024).

Both in tourist and institutional discourse, anthroponyms and toponyms seem to be the least subject to translation (Corpas Pastor & Sánchez Rodas 2022). However, the frequency of LOC entities is much higher in German > Spanish tourism. PER and ORG, which were representative in EU institutional texts, are the least present categories of tourist discourse (Corpas Pastor & Sánchez Rodas 2023). Full translations of names are less frequent in the tourism corpus, but mediated and non-mediated schemes are more varied than in institutional names (Sánchez Rodas 2023). This result, which may seem paradoxical, is the consequence of a more intensive use of paralinguistic signs in tourist discourse. Whereas EU texts occasionally resort to parentheses to encapsulate extra translations or borrowings, German and Spanish tourist texts consistently employ different types of inverted commas to draw the reader's attention to specific names, translated or not. This could also be related to the more unrestrained nature of tourist discourse, designed first to appeal and then to inform.

In a modern communication landscape which is increasingly multi-lingual, multimodal and syncretic, names are expected to gain relevance, thanks to their potential universality, their harmony with non-verbal semiotics, and their communicative economy. Professional translators would do well in systematizing the ins and outs of their (re)construction in the target language(s), where small clarifications or brackets can make a big difference. A closely related question is the current capacity of machine translation and AI translation to perform this clockwork task with enough quality in a sufficient combination of languages. Human translators and linguistic machine trainers have much to contribute to this possible advancement. The fine-tuning between semantic specificity and accessibility often needed in name translation and the wide range of pragmatic possibilities may be still hard to perform by automated systems on their own.

All our studies on name translation and interpreting so far revealed their capacity to unveil the innermost motivations of language production and transfer thanks to their subtlety. Names are highly referential micro-units of understanding, capsules of what is most physical in languages: concrete places, unrepeated human subjects, one-time historical events, everyday products. This also makes them a focus of conflict and attention, as physical representations unleash human emotions and power dynamics. The relevance of names in society still makes rather necessary to study their translation and interpreting, as well as their close links to paralinguistic and non-linguistic communication in the wider framework of multimodal semiotics and AI advancement.

## References

- ALLERTON, David J. (1987) "The linguistic and sociolinguistic status of proper names: what are they, and who do they belong to?" *Journal of Pragmatics* 11:1, pp. 61-92.
- CORPAS PASTOR, Gloria & Fernando Sánchez Rodas. (2022) "NLP-enhanced shift analysis of named entities in an English < > Spanish intermodal corpus of European petitions." In: Kajzer-Wietrzny, Marta *et alii* (eds.) 2022. *Mediated discourse at the European Parliament*. Berlin: Language Science Press, pp. 219-251.
- CORPAS PASTOR, Gloria & Fernando Sánchez Rodas. (2023) "EU phraseological verbal patterns in the PETIMOD 2.0 corpus: a NER-enhanced approach." In: Biel, Łucja & Hendrik J. Kockaert (eds.) 2023. *Handbook of Legal Terminology*, vol. 3. Amsterdam & Philadelphia: Benjamins, pp. 397-430.
- CUÉLLAR LÁZARO, Carmen. (2016) "Proper names in audiovisual translation: dubbing vs. subtitling." *Namenkundliche Informationen* 107/108, pp. 117-134. <https://doi.org/10.58938/ni551>
- DEL ÁGUILA, Eva. (1997) *La traducción de nombres propios en el ámbito jurídico y económico*. Unpublished PhD thesis. Granada: Universidad de Granada.
- DURÁN MUÑOZ, Isabel. (2011) "Tourist translations as a mediation tool: misunderstandings and difficulties." *Cadernos de Tradução* 27:1, pp. 24-49. <http://www.periodicos.ufsc.br/index.php/traducao/article/view/2175-7968.2011v1n27p29>



- DURÁN MUÑOZ, Isabel. (2012a) "Caracterización de la traducción turística: problemas, dificultades y posibles soluciones." *Revista de Lingüística y Lenguas Aplicadas* 7:1, pp. 104-113. +<http://ojs.cc.upv.es/index.php/rdlyla/article/view/1127>
- DURÁN MUÑOZ, Isabel. (2012b) "Analysing common mistakes in translations of tourist texts (Spanish, English and German)." *Onomázein* 26:2, pp. 335-349.
- FRANCO AIXELÁ, Javier. (2000) *La traducción condicionada de los nombres propios (inglés-español)*. Salamanca: Almar.
- FUENTES LUQUE, Adrián. (2016) "Branding and selling a country through translated tourism advertising: Spain's image." *Revista de lenguas para fines específicos* 22:2, pp. 84-103. <http://dx.doi.org/10.20420/rife.2016.113>
- FUENTES LUQUE, Adrián & Alexandra Santamaría Urbieto. (2020) "Machine translation systems and guidebooks: an approach to the importance of the role of the human translator." *Onomázein* 7, pp. 63-82. <http://ojs.uc.cl/index.php/onom/article/view/29659>
- GAŁKOWSKI, Artur. (2017) "Chrematonimia w kulturze współczesnej." *Onomastica* 61:1, pp. 55-71. <https://onomastica.ijp.pan.pl/index.php/ONOM/article/view/132>
- HERMANS, Theo. (1988) "On translating proper names, with reference to De Witte and Max Havelaar." In: Wintle, Michael (ed.) 1988. *Modern Dutch Studies*. London: The Athlone Press, pp. 11-24.
- JACQUET, Guillaume *et al.* (2019) "Cross-lingual linking of multi-word entities and language-dependent learning of multi-word entity patterns." In: Parmentier, Yannick & Jakub Waszczuk (eds.) 2019. *Representation and parsing of multiword expressions*. Berlin: Language Science Press, pp. 269-298.
- KELLY, Dorothy. (2005) "'Lest Periko Ortega give you a sweet ride...' o la urgente necesidad de profesionalizar la traducción en el sector turístico." In: Fuentes Luque, Adrián (ed.) 2005. *La traducción en el sector turístico*. Granada: Atrio, pp. 155-170.
- MAŃCZAK, Witold. (1969) "Différence entre nom propre et nom commun." In: Hornung, Herwig H. (éd.) 1969. *Abhandlungen des X. Kongresses für Namenforschung*, vol. 2. Wien: Akademischer Verlag, pp. 285-291.
- MARINA, María Cruz. (2019) "Baden-Wurtemberg, la región donde brillan todos los colores de Alemania." <https://www.eleconomista.es/turismo-viajes/>

- noticias/9825167/04/19/BadenWurtemberg-la-region-donde-brillan-todos-los-colores-de-Alemania.html
- MILL, John Stuart. (1875) *A System of Logic, ratiocinative and inductive*. London: Longmans, Green, Reader, and Dyer.
- MINKOV, Michael. (2013) "The concept of culture." In: Minkov, Michael (ed.) 2013. *Cross-Cultural Analysis*. Los Angeles: Sage, pp. 9-19.
- MOYA, Virgilio. (2000) *La traducción de los nombres propios*. Madrid: Cátedra.
- NEWMARK, Peter. (1988) *A Textbook of Translation*. Essex: Longman/Pearson.
- NOBS FEDERER, Marie-Louise. (1995) "Reflexiones sobre la problemática específica de la traducción de textos procedentes del campo del turismo." In: Iniesta Mena, Eva (ed.) 1995. *Perspectivas hispanas y rusas sobre la traducción*. Granada: Método Ediciones, pp. 239-259.
- POLYAKOVA, Oksana. (2021) "Turismo y traducción de guías de viajes en línea ruso-español." In: Candel Mora, Miguel Ángel (ed.) 2021. *Lenguas de especialidad, turismo y traducción*. Valencia: Tirant Humanidades, pp. 69-78.
- RAMÍREZ ALMANSA, Isidoro. (2019) "La traducción (español-alemán) de los referentes culturales en la película *Un franco, 14 pesetas*." In: Alcalde Peñalver, Elena et alii (eds.) 2019. *El sector de la traducción. La perspectiva del traductor novel*. Alcalá de Henares: Universidad de Alcalá, pp. 91-115.
- SÁNCHEZ RODAS, Fernando. (2023) "La traducción e interpretación de nombres de organizaciones en el eurolecto inglés y español: un estudio basado en corpus." *Romanica Olomucensia* 35:1, pp. 157-172. DOI: 10.5507/ro.2023.011
- SÁNCHEZ RODAS, Fernando. (2024) *English < > Spanish Translation and Interpreting of Proper Names in the European Union Discourse*. Málaga: UMA Editorial. Online version: <https://hdl.handle.net/10630/30929>
- SÁNCHEZ RODAS, Fernando & Raquel Del Río López. (2024) "Multimodal translation in YouTube shorts from Spanish football teams: Real Madrid against Barcelona." *Transletters* 8:1, pp. 1-33. <https://journals.uco.es/tl/article/view/17277/15920>
- SÄRKKÄ, Heikki. (2007) "Translation of proper names in non-fiction texts." *Translation Journal* 11:1. <https://translationjournal.net/journal/39proper.htm>
- SEVERAL AUTHORS (Duden). (2014) *Rechtschreibung und Grammatik*. Berlin: Dudenverlag.

- SEVERAL AUTHORS (Tourismus-Marketing GmbH Baden-Württemberg). (2024) "Traumurlaub im Süden Deutschlands: Urlaubsland Baden-Württemberg." <https://www.tourismus-bw.de/>
- SOTO ALMELA, Jorge. 2013. "La traducción de términos culturales en el contexto turístico español-inglés: recepción real en usuarios anglófonos." *Quaderns* 20, pp. 235-250.
- VENCESLÁ MARISCAL, Isabel. 2005. "¿Peña o 'flamenco association'? sobre las dificultades de la traducción de términos del flamenco." In: Fuentes Luque, Adrián (ed.) 2005. *La traducción en el sector turístico*. Granada: Atrio, pp. 139-144.

## BIONOTE / NOTA BIOGRÁFICA

FERNANDO SÁNCHEZ RODAS is PhD in Linguistics, Literature and Translation by the University of Málaga (UMA), with a thesis on the English < > Spanish translation and interpreting of proper names at the EU Parliament. He also holds one MA in Audiovisual Translation (English/German > Spanish) and one MA in Legal and Financial Translation (English/French/German > Spanish). Fernando taught language didactics and English and German translation courses at UMA for more than seven years. His lines of research include multilingual Onomastics, parliamentary discourse, AI translation assessment, and new forms of audiovisual translation.

FERNANDO SÁNCHEZ RODAS es Doctor en Lingüística, Literatura y Traducción por la Universidad de Málaga (UMA) con una tesis sobre la traducción e interpretación de nombres propios inglés < > español en el Parlamento Europeo. También posee un Máster en Traducción Audiovisual (Inglés/Alemán > Español) por la Universidad de Cádiz y un Máster en Traducción Jurídica y Financiera (Inglés/Alemán/Francés > Español) por la Universidad de Córdoba. Fernando ha impartido asignaturas de didáctica y de traducción de inglés y alemán en la UMA durante más de siete años. Sus líneas de investigación incluyen la onomástica multilingüe, el discurso parlamentario, la evaluación de la traducción generada por IA y las nuevas formas de traducción audiovisual.