

FLORENT THOUVENIN* / DANIEL GERBER** / TILMANN ALTWICKER***

Trademark Opposition Proceedings in Switzerland: An Empirical Study of Legal Reasoning

This study is the first empirical analysis of legal reasoning in trademark opposition proceedings in Switzerland. We examine a novel dataset on trademark opposition proceedings brought before the Swiss Federal Institute of Intellectual Property (IPI). In these proceedings, the likelihood of confusion between two (or more) trademarks is assessed based on the similarity of the trademark signs and the similarity of the goods and services, taking into account a series of additional aspects such as the distinctive character of the opposing trademark and the level of attention of the average consumer when buying the goods and services for which the earlier trademark is registered. Our dataset contains information on 2,453 cases relating to proceedings between June 2002 and August 2018. In particular, we examine which substantive factors drive the outcome of these decisions. Some of our findings call into question the established legal doctrine. For example, our data suggest that the importance of the beginning of words for establishing similarity between word marks is overrated by legal doctrine. Furthermore, our data show no clear influence of the level of attention on the assessment of the likelihood of confusion. Instead, we found striking differences between the success rates of different types of trademarks. In fact, the data reveal a sliding scale with word marks being the most successful trademarks followed by figurative trademarks that contain a word element, and purely figurative trademarks. Based on our empirical findings, we make suggestions on how to improve the legal reasoning when assessing the likelihood of confusion.

I. Introduction

One of the salient issues in trademark law is the assessment of the likelihood of confusion between two trademarks. While the substantive law on this issue is settled, there is an almost complete lack of empirical studies on the legal reasoning involved.¹ This is unfortunate, because the impact of some of the factors taken into consideration by courts and trademark offices – e.g. the level of distinctiveness of a trademark or the level of attention of the

average consumer – remains unclear in legal doctrine. Unlike all other areas of intellectual property (IP) law, there is an impressive number of decisions dealing with the likelihood of confusion in trademark law. While courts address the issue on a regular basis, the overwhelming majority of decisions are taken by trademark offices. Accordingly, the empirical analysis conducted in this study is based on the examination of a large dataset on trademark opposition proceedings brought before a national trademark office, namely the Swiss Federal Institute of Intellectual Property (IPI).

With regard to the assessment of the likelihood of confusion as discussed in this paper, Swiss law is basically identical with European trademark law as regulated in the EU Trademark Regulation (TMR)² and the harmonized national trademark laws of all EU Member States.³ Hence, the findings of this study should be of interest throughout Europe. Despite some notable differences

* Professor for Information and Communications Law, University of Zurich, Faculty of Law.

** M.Sc. in Mathematics, University of Zurich, Faculty of Law.

*** Swiss National Science Foundation (SNSF) Professor for Public Law, International Law, Legal Philosophy and Empirical Legal Research, University of Zurich, Faculty of Law.

¹ While there are a few empirical studies on US and Australian trademark law (namely Barton Beebe, ‘An Empirical Study of the Multifactor Tests of Trademark Infringement’ (2006) 94 California Law Review 1581; Kevin Blum and others, ‘Consistency of Confusion? A Fifteen-Year Revisiting of Barton Beebe’s Empirical Analysis of Multifactor Tests of Trademark Infringement’ [2010] Stanford Technology Law Review 3; Vicky Huang, Kimberlee Weatherall and Elizabeth Webster, ‘The use of survey evidence in Australian trademark and passing off cases’ in Andrew Kenyon, Ng-Loy Wee Loon and Megan Richardson (eds), *The law of reputation and brands in the Asia Pacific* (Cambridge University Press 2012) and one study on EU trademark law (Ilanah Fhima and Catrina Denvir, ‘An Empirical Analysis of the Likelihood of Confusion Factors in European Trade Mark Law’ (2015) 46 IIC 310), there is no such study analyzing Swiss trademark law. Most importantly, the studies on US, Australian and EU trademark law only analyze a limited number of decisions rendered by courts and not by trademark offices (Barton Beebe analyzed 331 opinions of US Federal District Courts, Kevin Blum and others analyzed 206 decisions of the US District Court for the Southern District of New York, Vicky Huang, Kimberlee Weatherall and Elizabeth Webster analyzed 78 decisions of Australian courts, Ilanah Fhima and Catrina Denvir 2015 analyzed 106 decisions of the EU General Court).

² Regulation (EU) 2017/1001 of the European Parliament and of the Council of 14 June 2017 on the European Union trademark (codification) [2017] OJ L154/1. This regulation is supplemented by Commission Delegated Regulation (EU) 2018/625 of 5 March 2018 supplementing Regulation (EU) 2017/1001 of the European Parliament and of the Council on the European Union trademark, a repealing Delegated Regulation (EU) 2017/1430 [2018] OJ L104/1.

³ Directive (EU) 2015/2436 of the European Parliament and of the Council of 16 December 2015 to approximate the laws of the Member States relating to trademarks (Recast) [2015] OJ L336/1. No longer in force: First Council Directive 89/104/EEC of 21 December 1988 to approximate the laws of the Member States relating to trade marks [1989] OJ L40/1; Directive 2008/95/EC of the European Parliament and of the Council of 22 October 2008 to approximate the laws of the member States relating to trade marks (codified version) [2008] OJ L299/25.

between Swiss, European and US trademark law,⁴ similar considerations are used for assessing the likelihood of confusion in all of these jurisdictions.⁵ Therefore, at least some of our findings should also be of interest from a US perspective.

There are a number of factors that are relevant when assessing the likelihood of confusion between two conflicting trademarks. These factors include the similarity of the goods and services, the similarity of the trademark signs, the level of attention of the average consumer and the distinctive character of the trademark.⁶ Although there is a vast judicial practice of the IPI and the Swiss courts on the interpretation of each of these factors, their impact on actual decision-making by the IPI when assessing the likelihood of confusion is unknown.⁷ Therefore, this paper aims to empirically test the impact of some of these factors on the assessment of the likelihood of confusion by the IPI.

This article uses a unique dataset that allows for the empirical analysis of the legal determinants of the IPI's decision-making in trademark opposition proceedings. Our data include information on all trademark opposition proceedings brought before the IPI between June 2002 and August 2018. The dataset comprises 2,453 decisions by the IPI.

The article proceeds as follows: Part II describes the legal and institutional background of trademark opposition proceedings in Europe, Switzerland and the US. Part III contains some general remarks on the object and aim of empirical legal research, Part IV sets out the dataset, the variables and the hypotheses regarding the legal reasoning in trademark opposition proceedings. Part V outlines the results. The final Part VI briefly concludes by discussing some implications from our key findings.

II. Legal and institutional background

The most important issue in trademark law is the assessment of the likelihood of confusion. Under European trademark law and the trademark laws of all Member States of the European Union, Switzerland and the US, the likelihood of confusion is evaluated taking into consideration the identity or similarity of the trademark signs and the identity or similarity of the goods or services covered by the trademarks (Art. 8(1)(b) TMR; Art. 5(1)(b) TMD; Art. 3(1)(a) Swiss Trademark Protection Act (TMA);⁸ 15 USC § 1052(d) Trademark Act of 1946).⁹

⁴ eg the existence of common law trademarks (marks which are in use but not registered, 15 USC § 1127 (2)) or the *ex officio* examination of confusing similarity with trademarks already registered or applied for in the course of the registration procedure. (15 USC § 1052(d) Trademark Act of 1946). Trademark Manual of Examining Procedure (TMPEP), United States Patent and Trademark Office, October 2018, 1207.01 <<https://tmep.uspto.gov/RDMS/TMEP/current#/current/TMEP-1200d1e5044.html>> accessed 4 May 2021. See also David W Barnes, *Trademark and unfair Competition Law* (Edward Elgar Publishing 2014) 11 and 135 (obtaining trademark rights through use), 13 (search for conflicting trademarks); Graeme B Dinwoodie and Mark D Janis, *Trademark and Unfair Competition, Law and Policy* (Wolters Kluwer 2014) 247 (common law trademark), 343 (search for conflicting trademarks).

⁵ 15 USC § 1052(d); eg so called DuPont factors (see *In re E. I. DuPont De Nemours & Co.* 476 F.2d 1357 (CCPA 1973), 1361 ff; see also *In re Majestic Distilling Co.* 315 F.3d 1311 (Fed. Cir. 2003)); Trademark Trial and Appeal Board Manual of Procedure (TBMP), June 2019, 309.03 (c) (2) B, 300-78 ff; for an overview of factor tests in likelihood-of-confusion analysis see Dinwoodie and Janis (n 4) 519 ff.

⁶ See IPI guidelines for trademark issues, 1 January 2019, 256-268.

⁷ For an overview see Gallus Joller, 'Art. 3' in Michael Noth, Gregor Bühler and Florent Thouvenin (eds), *Markenschutzgesetz (MSchG)* (2nd edn, Stämpfli 2017) paras 73ff.

The two factors are first appraised separately and then in combination. In this second step, the interaction between the two factors becomes relevant, i.e. in the event of a high level of similarity (or even identity) with respect to the goods and services a relatively low level of similarity with regards to the trademark signs (and vice versa) is sufficient to lead to a likelihood of confusion.¹⁰

The likelihood of confusion is the key issue in both infringement proceedings brought before civil courts and in opposition proceedings brought before the European Intellectual Property Office (EUIPO), national trademark offices of EU Member States, the Trademark Trial and Appeal Board (TTAB) of the United States Patent and Trademark Office or – in our case – the IPI. In all of these proceedings, the earlier trademark, i.e. the trademark which has been filed (or – for the US – used) first, prevails if the court or trademark office finds a likelihood of confusion (Art. 8(1) TMR; Art. 5(1) TMD; Art. 3 (1) TMA; 15 USC § 1052(d) Trademark Act of 1946)). The likelihood of confusion is assessed with regard to the similarity of the trademark signs and the goods and services for which the trademarks are registered based on the overall impression of the trademarks¹¹ and taking into account

⁸ Bundesgesetz über den Schutz von Marken und Herkunftsangaben (*Markenschutzgesetz*, MSchG), 28 August 1992, SR 232.11; art 3 TMA reads as follows: '(1) Also excluded from trademark protection are signs that are: (a) identical to an earlier trademark and are intended for the same goods or services; (b) identical to an earlier trademark and intended for similar goods or services such that a likelihood of confusion results; (c) similar to an earlier trademark and intended for the same or similar goods or services such that a likelihood of confusion results. (2) An earlier trademark is: (a) a filed or registered trademark that gives rise to a right of priority under this Act (art 6-8); (b) a trademark that is well-known in Switzerland within the meaning of Article 6^{bis} of the Paris Convention for the Protection of Industrial Property of 20 March 1883 (Paris Convention) at the time of filing the sign referred to in paragraph 1. (3) The grounds for refusal under this Article may only be invoked by the proprietor of the earlier trademark.' [Translation adopted from Michael Noth, Gregor Bühler and Florent Thouvenin (eds), *Markenschutzgesetz (MSchG)* (2nd edn, Stämpfli 2017)].

⁹ For an overview of factor tests in likelihood-of-confusion analysis see Dinwoodie and Janis (n 4) 519 ff; as an example, see also *In re E. I. DuPont De Nemours & Co.* 476 F.2d 1357 (CCPA 1973), 1361f.

¹⁰ Swiss Federal Supreme Court, BGE 122 III 382, 387; Joller (n 7) para 47; Matthias Staedeli and Simone Brauchbar Birkhäuser, 'Art. 3' in Lucas David and Markus R Frick (eds), *Markenschutzgesetz (MSchG)* (3rd edn, Helbing Lichtenhahn 2017) para 154. For EU law: EUIPO, 'Trade mark guidelines', pt C, s 2, ch 7 Global Assessment, 2 Interdependence Principle; Case C-342/97 *Lloyd Schuhfabrik Meyer & Co GmbH v Klijsen Handel BV* ECLI:EU:C:1999:323, para 19; Case C-234/06 *P Il Ponte Finanziaria SpA v OHIM (Trade Marks and Designs)* ECLI:EU:C:2007:514, para 48; Case C-39/97 *Canon Kabushiki Kaisha v Metro-Goldwyn-Mayer Inc* ECLI:EU:C:1998:442, para 17; Andrea Jaeger-Lenz, 'Art. 8' in Gordian Hasselblatt (ed), *European Union Trade Mark Regulation* (2nd edn, Hart Publishing 2018) para 46; Stefan Hennigs, 'Art. 8' in Annette Kur, Verena v Bomhard and Friedrich Albrecht (eds), *Markenrecht* (2nd edn, CH Beck 2018) para 121; Günther Eisenführ and Ulrich Sander, 'Art. 8' in Günther Eisenführ and Detlef Schennen (eds), *Unionsmarkenverordnung* (5th edn, Carl Heymanns 2016) paras 43 and 142. For U.S. law: Mary LaFrance, *Understanding Trademark Law* (4th edn, Carolina Academic Press 2020) 152; Deborah Bouchoux, *Intellectual Property, The Law of Trademarks, Copyrights, Patents, and Trade Secrets* (5th edn, Cengage Learning 2018) 126.

¹¹ Swiss Federal Supreme Court, BGE 128 III 441, 446; Swiss Federal Supreme Court, BGE 126 III 315, 321; Joller (n 7) para 127; Staedeli and Brauchbar Birkhäuser (n 10) para 41; Eugen Marbach, *Markenrecht* (2nd edn, Helbing Lichtenhahn 2009) para 864 ff; for EU law: Case T-292/01 *Phillips-Van Heusen Corp v OHIM (Trade Marks and Designs)* ECLI:EU:T:2003:264, para 47; Case C-342/97 *Lloyd Schuhfabrik Meyer & Co GmbH v Klijsen Handel BV* ECLI:EU:C:1999:323, para 25; Case C-251/95 *SABEL BV v Puma AG, Rudolf Dassler Sport, Sabèl* ECLI:EU:C:1997:528, para 23; EUIPO, 'Trade mark guidelines', pt C, s 2, ch 4 Comparison of signs, 3.1 Introduction; Jaeger-Lenz (n 10) paras 99 and 118; Hennigs (n 10) para 71; Eisenführ and Sander (n 10) para 64. For US law: See, eg, *AutoZone, Inc v Tandy Corp* 373 F.3d 786,

the distinctive character of the earlier trademark¹² and the level of attention of the average consumer when buying the goods and services that the earlier trademark is registered for.¹³ In Swiss and EU law, the likelihood of confusion is a legal question, not an empirical one.¹⁴ As a consequence, the likelihood of confusion is not subject to proof (using, for example, consumer surveys)¹⁵ and the

competent courts and the IPI retain an important margin of discretion when deciding on this issue.

In Switzerland, trademark opposition proceedings are administered by the IPI. The law governing trademark opposition proceedings can be found in Art. 31ff TMA.¹⁶ The decision of the IPI can be appealed to the Swiss Federal Administrative Court (Art. 31 and Art. 33(e) Federal Administrative Court Act). The decision of this court is final, there being no appeal to the Swiss Federal Supreme Court (Art. 73 Federal Supreme Court Act). Instead of initiating opposition proceedings, the proprietor of an earlier trademark may initiate civil proceedings before the court of sole cantonal instance (Art. 5 Swiss Civil Procedure Code). Such proceedings may also be initiated if the opposition has been rejected by the IPI or the Swiss Federal Administrative Court. The decisions of civil courts on trademark infringements can subsequently be appealed to the Federal Supreme Court (Art. 72 Federal Supreme Court Act).

Trademark opposition proceedings allow the proprietor of an earlier trademark to react quickly after the registration of a later trademark which is deemed to be identical with or similar to the earlier trademark and has been registered for identical or similar goods or services. To do so, the proprietor of the earlier trademark must submit an opposition with the IPI against the later trademark within three months after the publication of the registration of the later trademark. By initiating opposition proceedings, the proprietor of the earlier trademark can avoid lengthy and costly civil proceedings and eliminate a later, confusingly similar trademark within a relatively short period of time and at relatively low cost. There are two parties involved: the *defendant*, who has recently filed a trademark, and the *opponent*, who claims that the newly registered trademark is confusingly similar to their earlier trademark. If the trademark office finds a likelihood of confusion, the later trademark will be deleted from the trademark register.

III. Object, aim and limitations of an empirical approach

Empirical legal research is about analyzing legal data.¹⁷ One can distinguish between a narrower and a wider concept of legal data. The narrow concept relates to computer-readable legal texts (such as court judgments or legal norms in HTML format). The wider concept of legal data encompasses any observable data that are relevant from a socio-legal, economic or other perspective on the law (for example, data on the political ideology of the adjudicators or the economic performance of a company). The present study remains within the narrower concept of legal data by analyzing the text of the decisions taken by the IPI in opposition proceedings.

785ff (6th Cir 2004); *Clicks Billiards, Inc v Sixshooters, Inc* 251 F.3d 1252, 1259 (9th Cir 2001); *The Forschner Group, Inc v Arrow Trading Co, Inc* 124 F.3d 402, 409 (2nd Cir 1997); *Union Carbide Corp v Eveready, Inc* 531 F.2d 366, 379 (7th Cir 1976); see also LaFrance (n 10) 147 ff; Bouchoux (n 10) 125.

¹² Joller (n 7) para 73; Staedeli and Brauchbar Birkhäuser (n 10) para 47; Marbach (n 11) para 975. For EU law: Case C-39/97 *Canon EU:C:1998:442*, para. 18 ff; EUIPO, 'Trade mark guidelines', pt C, s 2, ch 5 Distinctiveness of the earlier mark; Jaeger-Lenz (n 10) paras 173 ff; Hennigs (n 10) paras 54 ff; Eisenführ and Sander (n 10) paras 78 ff and 144 ff. For US law: See, eg, *Hormel Foods Corp v Jim Henson Prods* 73 F.3d 497, 502ff (2nd Cir 1996); *McGregor-Doniger Inc v Drizzle Inc* 599 F.2d 1126, 1131ff (2d Cir 1979); see also LaFrance (n 10) 153 ff; Bouchoux (n 10) 127.

¹³ Joller (n 7) para 50; Staedeli and Brauchbar Birkhäuser (n 10) para 164; Marbach (n 11) para 995. For EU law: Case C-498/07 P *Aceites del Sur-Coosur SA, formerly Aceites del Sur v European Union Intellectual Property Office* [2009] ECR I-7371, para 74; Case C-342/97 *Lloyd Schuhfabrik Meyer & Co GmbH v Klijsen Handel BV* ECLI:EU:C:1999:323, para 26; EUIPO, 'Trade mark guidelines', pt C, s 2, ch 3 Relevant public and degree of attention, 3 Defining the Degree of Attention; Jaeger-Lenz (n 10) paras 92 ff; Hennigs (n 10) paras 47 ff; Eisenführ and Sander (n 10) paras 47 ff. Under US law, there is a similar criterion referred to as 'consumer sophistication', see, eg, *Virgin Enters v Nawab*, 335 F.3d 141, 151 (2nd Cir 2003); *TCPIP Holding Co v Haar Communications, Inc* 244 F.3d 88, 102 (2nd Cir 2003); see also LaFrance (n 10) 156 ff; Bouchoux (n 10) 127.

¹⁴ Swiss Federal Supreme Court, BGE 128 III 96, 97; Swiss Federal Supreme Court, BGE 126 III 315, 317; Joller (n 7) para 39; Marbach (n 11) para 942; Staedeli and Brauchbar Birkhäuser (n 10) para 32. For EU law: EUIPO, 'Trade mark guidelines', pt C, s 2, ch 1 General principles, 3.4 Likelihood of confusion: questions of fact and questions of law; Jaeger-Lenz (n 10) para 50; Eisenführ and Sander (n 10) para 46. In the US, a majority of courts treat the assessment of the likelihood of confusion as a question of fact, see, eg, *Facenda v N.F.L. Films, Inc* 542 F.3d 1007, 1024 (3rd Cir 2008); *General Motors Corp v Urban Gorilla, LLC* 500 F.3d 1222, 1227 (10th Cir 2007); *Everest Capital, Ltd v Everest Funds Ltd, L.L.C.* 393 F.3d 755, 760 (8th Cir 2005). However, some courts treat it as a legal question, see, eg, *Starbucks Corp v Wolfe's Borough Coffee, Inc*, 588 F.3d 97, 105 (2nd Cir 2009); *Leelanau Wine Cellars v Black & Red, Inc* 502 F.3d 504, 515 (6th Cir 2007); *Star Indus. v Bacardi & Co., Ltd* 412 F.3d 373, 384 (2nd Cir 2005); *China Healthways Inst, Inc, v Wang* 491 F.3d 1337, 1339 (Fed Cir. 2007). A recent Supreme Court decision now seems to support the majority, see *Hana Financial, Inc v Hana Bank* 135 S. Ct. 907, 911 (2015); however, this decision was related to so-called 'trademark tacking'. See for US Law also LaFrance (n 10) 145 ff; Dinwoodie and Janis (n 4) 524.

¹⁵ For Swiss Law: Swiss Federal Supreme Court, BGE 126 III 315, 317, stating that likelihood of confusion does not require proof of actual confusion and that the proof of actual confusion does not necessarily establish a legally relevant likelihood of confusion; see also Joller (n 7) para 42; Staedeli and Brauchbar Birkhäuser (n 10) para 32; Joller (n 7) para 42. For EU law: Case T-346/04 *Sadas SA v OHIM (Trade Marks and Designs)* ECLI:EU:T:2005:420, para 69, according to which 'it is not necessary to establish the existence of actual confusion, but the existence of a likelihood of confusion'; see also Hennigs (n 10) paras 24 and 130 ff. However, according to the EUIPO, incidences of actual confusion may weigh in favor of likelihood of confusion, see EUIPO, 'Trade mark guidelines', pt C, s 2, ch 6 Other factors, 4 Incidences of Actual Confusion. This is different under US law, see, eg, *Amstar Corp. v Domino's Pizza, Inc* 615 F.2d 252, 263 (5th Cir 1980): 'Although evidence of actual confusion is not necessary to a finding of likelihood of confusion, it is nevertheless the best evidence of likelihood of confusion.'; *International Kennel Club v Mighty Star, Inc* 846 F.2d 1079, 1089ff (7th Cir 1988); see also Bouchoux (n 10) 127, according to whom, although the occurrence of actual confusion is not required to establish likelihood of confusion, evidence of confusion by consumer testimony or survey evidence is nevertheless highly persuasive with regard to the establishment of likelihood of confusion; LaFrance (n 10) 143, with further references. For a detailed discussion of the impact and desirability of consumer surveys in the assessment of likelihood of confusion see Huang, Weatherall and Webster (n 1); for a recent empirical study on the use and impact of consumer surveys in trademark litigation see Katie Brown, Natasha Brison and Paul Batista, 'An Empirical Examination of Consumer Survey Use in

Trademark Litigation' (2019) 39 *Loyola of Los Angeles Entertainment Law Review* 237.

¹⁶ art 31 TMA: (1) The proprietor of an earlier trademark may file an opposition to a registration on the basis of art 3 para 1. (1^{bis}) He may not file an opposition to a registration of a geographical mark. (2) The opposition must be submitted in writing to the IPI with a statement of reasons within three months of publication of the registration. The opposition fee must also be paid within this time limit. [Translation adopted from Joller (n 7)].

¹⁷ See Tilmann Altwicker, 'International Legal Scholarship and the Challenge of Digitalization' (2019) 18 *Chinese Journal of International Law* 220, 221.

Three aims of empirical legal research can broadly be distinguished: First, in its socio-legal variant, empirical legal research is usually motivated by the aim of better understanding the functioning of courts and other public authorities as distinct social and political processes.¹⁸ For example, it is common to study the association between political ideology and outcomes of legal proceedings. Other studies examine the role of gender in legal decision-making. Second, policy-oriented empirical legal research is conducted with the aim of uncovering the potential for reform (e.g. enhancing the efficiency of proceedings).¹⁹ Third, some empirical legal research examines more directly the legal reasoning with the aim of uncovering divergences between the ‘law in action’ and the ‘law in the books’ or to find patterns in the legal reasoning that are unknown to legal doctrine. This strand of empirical legal research is especially helpful in narrowing down the set of plausible doctrinal explanations of the law.²⁰ The present study falls into the last category. Its main goal is to test legal doctrine in trademark law empirically. In this way, we are able to show where legal practice conforms to legal doctrine, but also where it diverges. In our view, the aim of empirical legal research is not to test if legal doctrine is applied properly. Quite to the contrary, observable divergences between ‘law in action’ and ‘law in the books’ are taken as an indicator that legal doctrine must be critically re-assessed and possibly adapted to ensure that it can serve its purpose of explaining legal reasoning in past decisions and serving as guidance in future cases.

While empirical legal research offers the benefit of a methodologically rigorous approach, it is, of course, not without limitations with respect to explaining the ‘law in action’. First, some legal concepts (and often the most interesting ones) are difficult or even impossible to measure (for example, it is often difficult to capture the ‘gravity’ of a violation of legal rights). Some legal concepts have a different meaning depending on the context, or their use is not standardized, making an empirical measurement difficult. In some cases, measurement is impossible because of the unlimited variety of concepts that are used in practice. This is problematic because, whereas legal analysis often deals with individual cases, empirical analysis is concerned with aggregate effects.²¹ To carry out a meaningful empirical analysis, the data must allow for some generalizations to be made. Sometimes, however, aggregation is impossible due to the variety observed in practice. For example, in our case, a variable capturing the similarity of goods and services – although of key importance from a legal perspective – could not be captured due to the virtually unlimited variety of goods and services that can be claimed and the lack of standardization of the language used to describe them. Furthermore, some legally important aspects are hard to extract from legal texts and thus not amenable to empirical analysis. For example, for the present study we decided not to code

variables capturing the sound and the meaning of word marks as we would have needed additional human coders, time and funding for this task although – from a legal perspective – this information would have been valuable. Second, the quality of empirical legal research depends to some degree on the number of observations that are analyzed. For example, our analysis of figurative marks is limited by the fact that our dataset contains only few cases relating to this type of trademark. We, therefore, focus mainly on word marks. Despite these general limitations, we believe that our analysis shows that empirical legal research provides important insights into the actual legal reasoning in trademark opposition proceedings that can and should inform the further development of legal doctrine.

IV. Data, variables and hypotheses

1. Dataset and unit of analysis

This study introduces a novel dataset that includes all decisions taken in trademark opposition proceedings brought before the IPI between June 2002 and August 2018 and decided between November 2005 and December 2018.²² The dataset was constructed semi-automatically using the data provided by the IPI, originally amounting to 2,665 documents. Prior to coding, decisions in Italian²³ and procedural decisions²⁴ were excluded, leading to a total of 2,610 documents.

If one or both parties engaging in opposition proceedings base their claims on more than one trademark, the IPI labels these cases as two or more separate opposition proceedings and issues two or more identical documents containing the decisions for all of these opposition proceedings.²⁵ Since no new information comes from the multiple coding of identical documents, these exact duplicates were removed, leading to a total of 2,460 documents. Furthermore, several documents were excluded because of improper scans or missing sections. In total, our dataset was constructed by coding 2,453 decisions.

In our dataset, 28.5% of the decisions were written in French and 71.5% in German. The earliest opposition proceedings recorded started in June 2002 and the last ended in August 2018. The median duration of proceedings is 315 days with the shortest being 31 days and the longest 9.3 years. More than 90% of all decisions were taken within two years after the filing of the opposition.

The 2,453 decisions were first analyzed by an algorithm designed to extract the variables automatically using regular expressions.²⁶ The encodings of this algorithm were then manually double checked and, if necessary, corrected by three human coders (law students). To aid inter-

¹⁸ Sharyn R Anleu and Kathy Mack, ‘Trial Courts and Adjudication’ in Peter Cane and Herbert Kritzer (eds), *The Oxford Handbook of Empirical Legal Research* (Oxford University Press 2010) 567.

¹⁹ *ibid.*

²⁰ See, *mutatis mutandis*, Holger Spamann, ‘Empirical Comparative Law’ (2015) 11 Annual Review of Law and Social Science 131, 138.

²¹ See Robert M Lawless and others, *Empirical Methods in Law* (2nd edn, Wolters Kluwer 2016) 10.

²² The first opposition that lead to a decision included in the dataset was lodged on 27 June 2002, the last on 18 August 2018; the first decision in our dataset was rendered by the IPI on 9 November 2005, the last on 19 December 2018.

²³ Decisions in Italian were excluded to reduce the complexity of the coding. This does not create a bias in our dataset given the low number of cases (41).

²⁴ For example, procedural decisions about the termination of opposition proceedings due to a settlement or a full or partial deletion of the defending trademark during the opposition proceedings. See for example, Decision by the IPI in opposition proceedings No 10049.

²⁵ See for example, Decision by the IPI in opposition proceedings No 9942 and No 9992, both *Tech Line Electronica v TECHLINE*.

²⁶ This was done using the software R <<https://cran.r-project.org/>> accessed 5 May 2021.

rather reliability across the two decision languages French and German, the human coders adhered to a codebook.²⁷

The majority of the decisions deal with the conflict between one opposing and one defending trademark. However, other constellations are possible, making it necessary to specify the unit of analysis. The IPI decides on the opposition of one opposing trademark against one defending trademark in a single opposition proceedings. If the opponent relies on or attacks several trademarks, several opposition proceedings are initiated and several decisions rendered. Opposition proceedings therefore always concern the encounter between two trademarks. Accordingly, the unit of analysis was chosen as being each encounter between one opposing and one defending trademark. This resulted in 2,603 encounters on which the IPI took a decision. In only 3.5% of all cases is there more than one opposing trademark and in only 1.3% of all cases is there more than one defending trademark. This implies that the choice of the unit of analysis will most likely not have a relevant impact on the findings. As a robustness check, the sample was also restricted to decisions which dealt with the encounter between one opposing and one defending trademark only. Thereafter, the findings did not change substantially.

2. Variables

We distinguish two groups of variables, one relating to the steps in the decision-making of the IPI and one relating to the substantive factors assessed within these steps.

a) Decision-making

A first group of variables relates to the data about the key steps in decision-making in opposition proceedings. This group includes a variable capturing the opponent's claim that its trademark is well-known in Switzerland,²⁸ a variable capturing the allegation of non-use of the opposing trademark by the defendant, a variable capturing whether the assessment of the similarity of goods and services by the IPI results directly in a rejection of the opposition, a variable capturing whether the assessment of the similarity of the trademark signs by the IPI results directly in a rejection and, finally, a variable capturing whether the assessment of the likelihood of confusion results in a rejection, or in partial or full approval of the opposition.

These variables allow us to identify the decisive issues in opposition proceedings and to highlight their relevance

²⁷ Intercoeder reliability was excellent: 246 cases (10% of all cases) were coded by all three coders. All but 6 variables were coded with a joint probability of agreement between 95% and 100%. The variable with the lowest interrater reliability was the 'level of attention' for which the coders disagreed in 30 cases (12.2% of all cases). This variable along with the 'level of distinctiveness' was reportedly the most difficult to code because of the sometimes inconsistent language used in the decisions.

²⁸ This relates to domestically non-registered but well-known trademarks (see art 3(2)(b) TMA). Whether the trademark is well-known in Switzerland is assessed on the basis of the circumstances of the individual case taking into account the following criteria: 'the notoriety of the trademark with the relevant consumers, the duration, scope and geographical extent of trademark use and trademark advertising, the duration and geographical extent of trademark registrations which have been made or applied for, the previous protection of the trademark right, in particular through recognition of notoriety by the competent authorities of individual contracting states, as well as the value associated with the trademark', see IPI guidelines for trademark issues, 1 January 2019, 239 (referencing the WIPO recommendation).

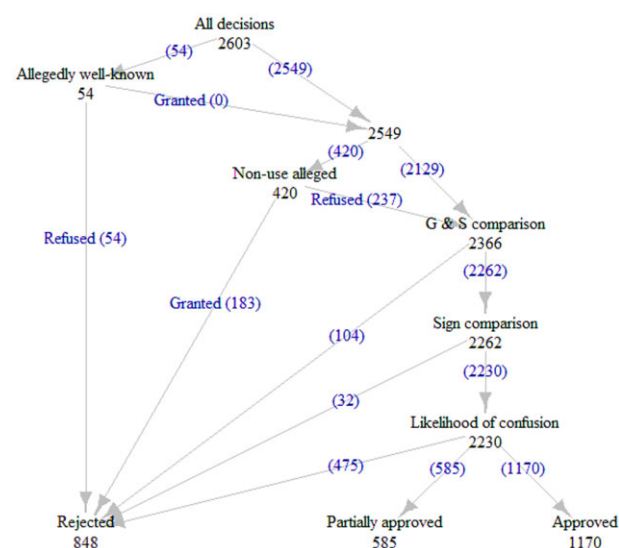


Figure 1. Decision-making procedure by the IPI for every encounter between an opposing and a defending trademark. For example, in 2,549 decisions non-use of the opposing trademark was claimed in 420 cases, upheld in 183 cases and rejected in 237 cases.

in a decision tree (see Figure 1).²⁹ In a first step, the IPI needs to check if the opposition is based on an earlier trademark. Most often, this is assessed by a simple comparison of the filing dates of the opposing trademarks. But an opposition can also rely on a trademark which has not been registered if the opponent claims that its trademark is well-known in Switzerland.³⁰ The first variable captures these rare cases, in which the opponent needs to allege and ultimately prove that its trademark is well-known. If the opponent fails to do so, the opposition is rejected without further examination. The second variable captures situations in which the defendant alleges a lack of use of the opposing trademark by claiming that this trademark has not been sufficiently used for at least five consecutive years. If the opponent fails to rebut this claim, the opposition is rejected without further examination.³¹

If these tests are passed, the next steps in the IPI's decision-making process are as follows: In a first step, the IPI assesses the similarity of the goods and services for which the conflicting trademarks are registered. If the trademarks are registered for dissimilar goods and services, the opposition is rejected. If not, the IPI continues its examination and assesses the similarity of the trademark signs. As with the similarity of goods and services, a complete lack of similarity between the trademark signs leads to a rejection of the opposition. If the IPI finds some similarity, it brings together its findings in the previous two steps to assess the likelihood of confusion in the ultimate step of its decision-making, which will result in a rejection or in a (full or partial) approval of the opposition. In the event of a partial approval, the opposition is only granted for a limited number of goods and services and rejected for the remainder.

²⁹ See below, Figure 1, IV.2.a.

³⁰ art 3(2)(b) TMA (Relative grounds for refusal).

³¹ art 12 TMA (Consequences of non-use).

In addition to these key steps of the decision-making process, data on the language³² of the decision, the date of filing of the opposition and the date of the decision were coded.

b) Substantive factors

A second group of variables concerns the most important factors considered in trademark opposition proceedings, namely the similarity of goods and services and the similarity of the trademark signs, as well as the level of attention of the average consumer and the level of distinctiveness of the opposing trademark.

The goods and services for which the two conflicting trademarks are registered are listed in the decisions of the IPI along with the relevant classes of goods and services according to the Nice classification scheme.³³ Since the goods and services are listed in a non-standardized way, we were unable to code and assess this factor on the level of the individual goods and services. Instead, we had to restrict our analysis to a variable capturing the classes of goods and services for which the conflicting trademarks were registered. While this grossly simplifies a much more nuanced assessment carried out by the IPI, even this simplified approach yielded some interesting results.³⁴ For a more accurate analysis of this key substantive factor, future research would have to find a suitable way to perform a full comparison of the complete lists of goods and services.

A trademark may consist of any sign, namely words, designs, designs containing text elements, colors, the shape of goods or of the packaging of goods, scents, etc. (Art. 1(2) TMA). In practice, however, only the first three types are regularly used and registered. Therefore, the variable ‘trademark type’ was classified in word marks, figurative marks, figurative marks containing word elements and others. If the trademark type was a word mark, the words were recorded and compared to assess the similarity of the trademark signs. The first comparison accounted for an identical starting letter; additional comparisons were made regarding the identity of the first two, three, and four letters and the same was done for the last letters of the trademark signs. We also recorded the typeface, but did not capture the sound and the meaning of the words with a specific variable.

To account for the impact of the level of attention of the average consumer on the assessment of the likelihood of confusion, the attention level was recorded in the categories average, high (including slightly above average) and low (including slightly below average). Unfortunately, the decisions do not deal with the attention level consistently. Sometimes several levels of attention are mentioned, and no concluding summary statement is given, in other cases the level of attention is not mentioned at all. In these cases, no level of attention could be recorded.

Lastly, the level of distinctiveness determines the scope of protection of a trademark, and the distinctive character of the opposing trademark should therefore have a relevant impact on decisions taken in opposition

proceedings.³⁵ To assess this, a variable capturing the level of distinctiveness and a variable capturing the scope of protection were coded with the levels high, average and low, or unknown if the distinctiveness or the scope of protection respectively was not mentioned.³⁶

3. Hypotheses

Based on the current state of legal doctrine in trademark law and the established practice in trademark opposition proceedings, we tested five hypotheses empirically. The first hypothesis (H 1) relates to the impact of alleging non-use of the opposing trademark on the outcome of opposition proceedings while the other four hypotheses all relate to aspects that concern the assessment of the likelihood of confusion, namely the impact of the level of attention (H 2) and the level of distinctiveness (H 3) on the rejection rate, the impact of first letter matches on the finding of similarity of trademark signs (H 4) and the impact of the trademark type on the finding of likelihood of confusion.

H 1: There is a higher rate of rejection in cases in which non-use of the opposing trademark is alleged even if the allegation is rejected

Trademarks are only protected against the use of confusingly similar trademarks if they are actually used for the claimed goods and services on the market. Once the grace period of five years after the date of registration has lapsed, third parties can claim non-use of a trademark and request that the trademark be deleted for all the goods and services for which it has not been used (Art. 12(1) TMA). In Switzerland, such request can be made in specific proceedings before the IPI (Art. 35a ff TMA), but also in civil proceedings and opposition proceedings. If the court or the IPI finds a lack of use, the trademark will be deleted from the trademark register, and it can no longer be enforced against any third party.

While an opposition is always (fully or partly) rejected if the defendant successfully claims non-use of the opponent’s trademark for some or all of the goods and services, we expect alleging non-use also to have an impact on the result of opposition proceedings even if the IPI comes to the conclusion that the opponent’s trademark has been sufficiently used. We expect such a correlation because the scope of protection of a trademark depends on the extent of its distinctive character, i.e. its level of distinctiveness,³⁷ which positively correlates with a trademark’s use on the market.

Trademarks are considered to have an inherently distinctive character if they consist of an original word, image or the like, e.g. Novartis. But they may also acquire distinctiveness through intensive and long-lasting use.³⁸ In both cases, using a trademark has a positive effect on the level of distinctiveness and hence on the scope of protection. Assuming that non-use is not regularly claimed by defendants in opposition proceedings if the opposing trademark has been used to a relevant degree, we expect

³² The decisions are written in a single language, either German or French. But the list of goods and services for which the trademarks are registered is sometimes provided in the other language.

³³ <<https://www.wipo.int/classifications/nice/en/>> accessed 7 May 2021; Swiss Federal Supreme Court, BGE 122 III 382, 387.

³⁴ See below, Figure 2, V.1.b.

³⁵ IPI guidelines for trademark issues, 1 January 2019, 265–8.

³⁶ The same rule of classification and therefore the same caveats apply as for the level of attention.

³⁷ See below, Table 1, V.1.b and Table 2, V.2.

³⁸ Swiss Federal Supreme Court, BGE 122 III 382, 387; Staedeli and Brauchbar Birkhäuser (n 10) para 51; Joller (n 7) paras 76 and 102.

that oppositions are more often rejected if non-use has been alleged.

H 2: The higher the level of attention, the higher the rejection rate

According to established judicial practice,³⁹ literature⁴⁰ and the IPI guidelines for trademark issues,⁴¹ the likelihood of confusion depends on the attention level of the average consumer of the goods and services. For example, it is assumed that consumers are likely to be less attentive when buying 'mass articles of daily use',⁴² while goods and services offered to experts in their fields are expected to be regarded with a higher level of attention.⁴³ Accordingly, the level of attention is deemed to be higher or lower for some goods and services than for others.⁴⁴ In the case of a lower level of attention, the threshold for the likelihood of confusion is more easily met, whereas in the case of a higher level of attention, relatively minor differences may be sufficient to exclude a likelihood of confusion. Therefore, we expect that a higher level of attention leads to a higher rejection rate in opposition proceedings.

H 3: The higher the level of distinctiveness of the opposing trademark, the lower the rejection rate

According to established judicial practice,⁴⁵ literature⁴⁶ and the IPI guidelines for trademark issues,⁴⁷ the scope of protection of a trademark depends on its distinctiveness. Trademarks with a highly distinctive character have a larger scope of protection than trademarks with an average or low level of distinctiveness.⁴⁸ Therefore, relatively minor differences should suffice to exclude a likelihood of confusion if the distinctive character of the opposing trademark is limited, while more important differences are needed if such a trademark is highly distinctive. As a consequence, trademarks with a high level of distinctiveness are expected to prevail more often in opposition proceedings than trademarks with an average or low level of distinctiveness.

H 4: Similarity of word marks is more often found if the first and last letters of the words match

For word marks, the similarity of the trademark signs is assessed based on the typeface, the sound and the meaning of the words.⁴⁹ A similarity with regard to any of

these three factors can suffice to find similarity of the trademark signs.⁵⁰ According to established judicial practice,⁵¹ literature⁵² and the IPI guidelines for trademark issues,⁵³ the beginning and the ending of the words are deemed to be of particular importance for a finding of similarity between word marks. In parts, these approaches are supported by findings in perceptual psychology. According to research in this domain, humans focus on the beginning of a word when reading it, but the last letter of a word seems to be almost equally important.⁵⁴ With regard to the established (but not fully consistent) approaches in judicial practice and legal doctrine and the partial support of these approaches by insights gained in perceptual psychology, we were particularly interested in analyzing the assessment of similarity of word marks in IPI opposition proceedings. Based on legal practice and doctrine, we expect the similarity of the first and last letters of the opposing and defending word marks to have an impact on the assessment of the similarity of the conflicting trademarks and the likelihood of confusion.

H 5: Word marks prevail more often than any other type of trademarks

The scope of protection of a trademark depends on several factors. In addition to the level of distinctiveness,⁵⁵ the complexity of a trademark, i.e. the number of elements a trademark consists of, has an impact, e.g. rather long word marks (such as BMW's slogan 'The ultimate driving machine') will only be similar to a limited number of other trademarks, simply because there are more elements which can be different than in the case of a collision of relatively short trademarks (e.g. boss vs. boks⁵⁶). The same holds true if a trademark consists of several elements as in the case of figurative trademarks containing word elements. Given that word marks and figurative marks containing word elements are far more often involved in opposition proceedings than any other type of trademarks, we expect that word marks prevail more often in opposition proceedings than any other type of trademarks.

³⁹ Swiss Federal Supreme Court, BGE 122 III 382, 387ff.

⁴⁰ Joller (n 7) para 50; Staedeli and Brauchbar Birkhäuser (n 10) paras 162 ff; Marbach (n 11) para 995.

⁴¹ IPI guidelines for trademark issues, 1 January 2019, 264.

⁴² Swiss Federal Supreme Court, BGE 122 III 382, 388; Joller (n 7) paras 56 and 63; Staedeli and Brauchbar Birkhäuser (n 10) paras 164 and 167; Marbach (n 11) para 995; IPI guidelines for trademark issues, 1 January 2019, 264.

⁴³ IPI guidelines for trademark issues, 1 January 2019, 264.

⁴⁴ Swiss Federal Supreme Court, BGE 122 III 382, 388; Joller (n 7) paras 54 ff; Staedeli and Brauchbar Birkhäuser (n 10) paras N 167 ff; IPI guidelines for trademark issues, 1 January 2019, 264.

⁴⁵ Swiss Federal Supreme Court, BGE 122 III 382, 385.

⁴⁶ Joller (n 7) para 73; Staedeli and Brauchbar Birkhäuser (n 10) paras 45 and 47; Marbach (n 11) para 976.

⁴⁷ IPI guidelines for trademark issues, 1 January 2019, 264.

⁴⁸ Swiss Federal Supreme Court, BGE 128 III 441, 447; Swiss Federal Supreme Court, BGE 122 III 382, 385; Joller (n 7) para 73; Staedeli and Brauchbar Birkhäuser (n 10) paras 45 and 52; Marbach (n 11) paras 976, 979 and 981.

⁴⁹ Swiss Federal Supreme Court, BGE 122 III 382, 388; Swiss Federal Supreme Court, BGE 121 III 377, 379; Joller (n 7) para 140; Staedeli and Brauchbar Birkhäuser (n 10) para 59; Marbach (n 11) para 872; IPI

guidelines for trademark issues, 1 January 2019, 260. Similarly, EU trademark law features a visual, a phonetic and a conceptual comparison of the trademarks, the latter relating to the semantic content of the trademarks; see EUIPO, 'Trade mark guidelines', pt C, s 2, ch 4 Comparison of Signs, 3.4 Comparison of signs.

⁵⁰ Swiss Federal Supreme Court, BGE 4A_44/2007, [2008] sic! 295, 298; Joller (n 7) para 187; Staedeli and Brauchbar Birkhäuser (n 11) para 59; Marbach (n 11) para 875; IPI guidelines for trademark issues, 1 January 2019, 260.

⁵¹ Swiss Federal Supreme Court, BGE 122 III 382, 388; Swiss Federal Administrative Court, BVGer B-6173/2018 [3.3]; Swiss Federal Supreme Court, BGE 78 II 379, 381; Swiss Federal Supreme Court, BGE 126 III 315, 321; Swiss Federal Administrative Court, BVGer B-2635/2008 [6.2.2]; Swiss Federal Administrative Court, BVGer B-6146/2007 [8]; Swiss Federal Administrative Court, BVGer B-1637/2015 [4.1].

⁵² Joller (n 7) para 147ff; Staedeli and Brauchbar Birkhäuser (n 10) para 65; Marbach (n 11) para 881.

⁵³ IPI guidelines for trademark issues, 1 January 2019, 261.

⁵⁴ Christian Scheier and Andreas Lubberger, 'Vom Angriff der Tatsachen auf die Erfahrungsgrundsätze, Anmerkungen zur Physiologie und Psychologie der Markenwahrnehmungen' (2014) 11-12 MarkenR – Markenrecht 453, 466.

⁵⁵ See above, IV.3; see below Table 1, V.1.b and Table 2, V.2.

⁵⁶ Swiss Federal Supreme Court, BGE 121 III 377.

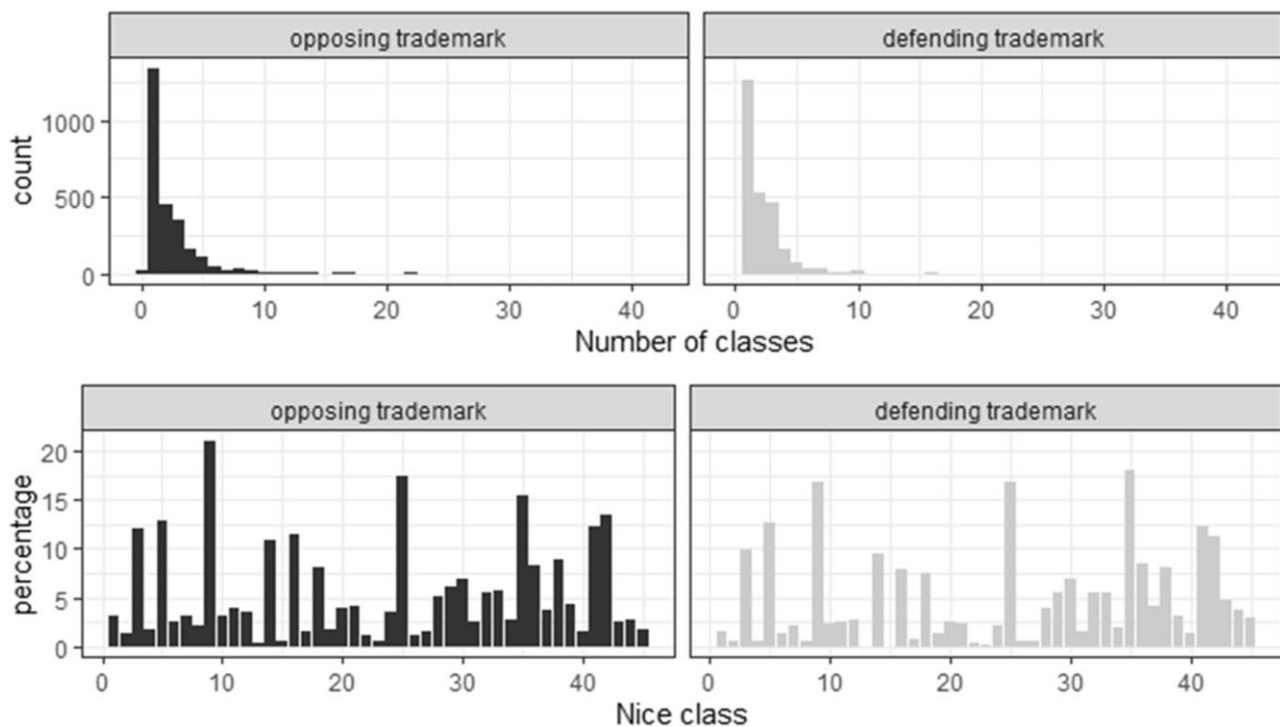


Figure 2. Top panel: Number of classes considered in the IPI's decision for the opposing trademark (left) or the defending trademark (right). Bottom panel: Share of opposing (left) or defending (right) trademarks per class. For example, for about 3% of opposing trademarks the opposition is (only or *inter alia*) based on goods and services of class 1 (chemical products for use in industry, science and agriculture).

V. Results

1. General findings

In addition to testing our hypotheses, the analysis of the data⁵⁷ also revealed some general findings that relate to the decision-making of the IPI and to some of the substantive factors.

a) Decision-making

The IPI decision-making steps are shown in Figure 1.

In 54 decisions, the opponent alleged that its trademark was well-known in Switzerland. In none of the cases did the IPI accept this allegation. The defendant claimed non-use of the opposing trademark in 420 cases and the IPI upheld non-use in 183 cases. The comparison of the goods and services resulted in 104 straight rejections, while the comparison of the trademark signs only led to a straight rejection in 32 cases. Accordingly, the likelihood of confusion was assessed in the vast majority of decisions, namely in 2,230 cases. Of these decisions, 21.3% (475) of the oppositions were rejected, 26.2% (585) were partially upheld and 52.5% (1170) were fully upheld.⁵⁸

Of all 2,603 decisions in our database, the IPI rejected 32.6% (848), partially upheld 22.5% (585) and fully upheld 44.9% (1170) of the oppositions. This shift to a higher rejection rate is caused by the straight rejection of an opposition in the case of non-use of the opposing trademark or a complete lack of similarity of goods and services or trademark signs.

b) Substantive factors

One of the two key substantive factors in the assessment of the likelihood of confusion is the similarity of the goods and services. Trademarks can be registered for an unlimited number of goods and services and, therefore, for as many classes as the Nice Classification contains. However, the histograms in the top panel of Figure 2 show that the majority of opposition proceedings deal with goods and services belonging to only one class. While opposition proceedings concerning goods and services from two or three classes are still common, only 15.2% of all opposition proceedings deal with goods and services from more than three classes. Note that some of the opposing trademarks were not registered for any goods and services because they had not been registered at all but were claimed to be well-known in Switzerland.⁵⁹

The bottom panel of Figure 2 shows which classes of goods and services are the subject of opposition proceedings.

Three classes account for more than 22% of all opposition proceedings: class 9 (mainly apparatus and

⁵⁷ All statistical analysis was conducted using R. <<https://cran.r-project.org/>> accessed 5 May 2021.

⁵⁸ At first sight, this finding appears to be quite different from that of Fhima and Denvir (n 1) 321, who reported that the EU General Court found a likelihood of confusion in 60% of the cases and no likelihood of confusion in 40% of the cases. However, the difference is merely a consequence of the different coding of the decisions. While Fhima and Denvir (n 1) coded each discrete trademark similarity analysis that the court conducted as a separate entry (Fhima and Denvir (n 1) 316, with the example that they coded the comparison of cosmetics and soap separately from the comparison of cosmetics and spa consultancy services because the General Court recorded separate findings in relation to each of these goods), we coded decisions which found likelihood of confusion for some but not for all the goods and services as partial approval. In fact, when

following the approach of Fhima and Denvir (n 1), the results are very similar.

⁵⁹ See above, IV.2.a.

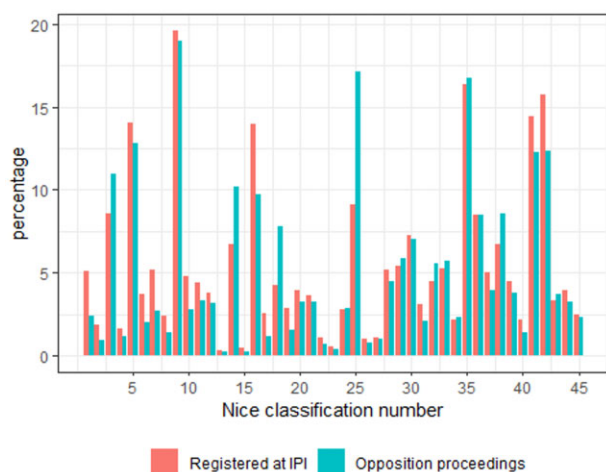


Figure 3. Average (median over the years 2002-2019) shares of trademarks per Nice class registered at the IPI, and shares of trademarks in opposition proceedings per Nice class in our dataset.

instruments for scientific or research purposes, audiovisual and information technology equipment, safety and life-saving equipment), class 25 (clothing, footwear, headgear), and class 35 (advertising, business management and administration, office functions). Conversely, some classes rarely occur: class 13 (firearms, fireworks), class 15 (musical instruments), and class 23 (yarns and threads for textile use) were only subject to one to three opposition proceedings in the 16 years captured by our dataset.

While these numbers correspond with the numbers of trademarks registered as shown in Figure 3, it is interesting that there are statistically significant divergences between the number of trademarks registered and the number of opposition proceedings in some classes.

Namely, there are only relatively few opposition proceedings in class 1 (mainly chemical products for use in industry, science, and agriculture) and class 16 (paper and cardboard, printed matter and certain goods made of these materials), while opposition proceedings are relatively frequent in class 3 (non-medicated cosmetics and toiletry preparations including cleaning preparations), class 14 (precious metals and certain goods made thereof, jewelry, clocks and watches), class 18 (leather, imitations of leather and goods made of these materials) and class 25 (clothing, footwear, headgear).

The other key substantive factor in the assessment of the likelihood of confusion is the similarity of the trademark signs, which is appraised differently for different types of trademarks. While there are well-established rules for assessing the similarity of word marks,⁶⁰ there is a lack of such rules for the encounter between other trademarks of the same type (e.g. figurative marks) and even more so for the encounter between trademarks of different types (e.g. figurative marks containing a word element vs. figurative marks). Accordingly, we were interested in knowing what trademark types were the subject of opposition proceedings how often and what encounters between similar and dissimilar trademark types occur.

The occurrence of the trademark types in our dataset is described in Figure 4. The left panel of Figure 4 shows that the opposing trademarks are predominantly word marks (68.2%), followed by figurative marks containing word elements (27.4%) and a remainder of figurative marks (4.3%) and only a few other types (0.1%). The defending trademarks, however, are split almost evenly between word marks (51.7%) and figurative marks containing word elements (46.1%). A small remainder (2.2%) consists of figurative marks, while no defending trademark was classified as other.

Contrary to the shares of Nice classes, the shares of trademark types represented in opposition proceedings are roughly equal to the shares of trademark types registered at the IPI.

The right panel of Figure 4 summarizes the encounters of trademark types. The most frequent encounter is between two word marks with 1,072 (41.2%) cases. The second most frequent encounter is an opposition by a word mark against a defending figurative mark containing word elements with 700 (26.9%) encounters. Encounters between figurative trademarks containing word elements account for 433 (16.6%) cases. The remaining possible combinations of encounters of trademark types account for only 398 (15.3%) of all cases.

In addition to the similarity of goods and services as well as trademark signs, the other most important factors for assessing the likelihood of confusion are the attention level of the average consumer and the level of distinctiveness of the opposing trademark, which defines its scope of protection. Given the importance of these substantive factors, we were interested in knowing how often and in which way they were assessed. The results are displayed in Table 1.

A level of attention was identified in only 39.7% of all encounters that reached the stage at which the likelihood of confusion was assessed (see Table 1). If a level of attention was identified, it was mostly high (47.1%) or average (38.1%) and, much less frequently, low (14.8%). A level of distinctiveness was identified in 68.3% of all cases reaching the assessment of the likelihood of confusion. A dominant share of 90.2% of all opposing trademarks was classified as having an average level of distinctiveness, whereas only 2.7% were classified as highly distinctive and only 7% as having a low level of distinctiveness. A scope of protection is mentioned in 73% of all cases in which the likelihood of confusion was assessed. Similar to the level of distinctiveness, the scope of protection is mostly average (93.7%), and rarely small (3.7%) or large (2.6%).

2. Alleging non-use

One might expect that non-use is alleged by the defendant in cases in which the opposing trademark has not been used at all or – at least – only to a limited extent.⁶¹ Accordingly, alleging non-use should be positively correlated with a rather low level of distinctiveness because a lack of or limited use of a trademark excludes the acquisition of distinctiveness and the low level of distinctiveness should lead to a higher rejection rate. This assumption led to our first hypothesis (H 1) according to which we

⁶⁰ See above, IV.3.

⁶¹ See above, IV.3.

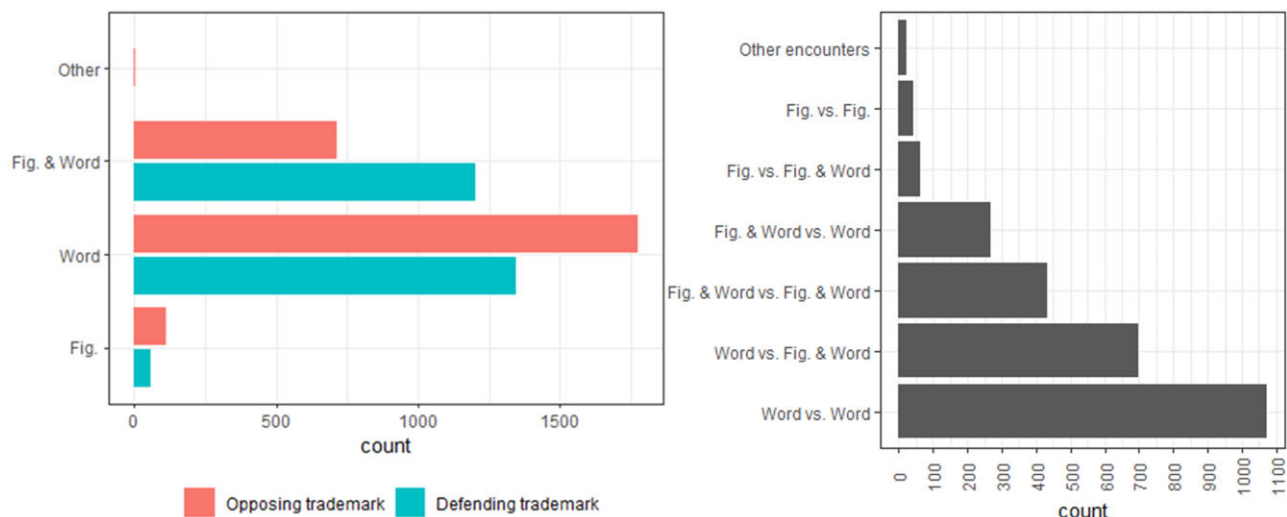


Figure 4. Left panel: Number of trademark types of the opposing and defending trademark. Right panel: Number of trademark type encounters, e.g. there are about 700 decisions relating to an opposing word trademark and a defending figurative trademark containing word elements.

expected a higher rejection rate in cases in which non-use of the trademark is alleged even if the allegation is rejected. In our data however, we found no evidence for a relationship between the alleged non-use of the opposing trademark and the level of distinctiveness.⁶²

Rather, our data show (see Table 2) that the distribution of the level of distinctiveness is quite similar in the cases in which non-use was not alleged and the cases in which non-use was alleged but rebutted. Somewhat surprisingly, the IPI found a high level of distinctiveness in 5.4% of the cases in which non-use was alleged but rebutted, as opposed to only 2.5% in cases in which non-use was not alleged. Given the small number of cases, however, this result is not robust.

The lack of a correlation between an alleged but rejected non-use and the level of distinctiveness can be explained by the fact that the IPI attributed an average level of distinctiveness in most of these cases. This indicates that non-use is often alleged with regard to trademarks which either have an inherently distinctive character or have been used to a normal extent leading to an average level of distinctiveness in most cases. This suggests that defendants often allege a lack of use of the opposing trademark for purely procedural reasons, namely in order to slow down the proceedings and extend the time until a decision is taken.

Alleging non-use does, however, increase the chance of rejection for the overall outcome of the proceedings from 27.7% (if non-use is not alleged) to 57.9% (if non-use is alleged). But this remarkable increase is only driven by the cases where non-use is rightly claimed and the opposition is rejected for lack of use of the opposing trademark.

After all, if non-use was alleged by the defendant but rejected by the IPI, the rate of rejection of the opposition decreased at the likelihood of confusion step from 16.9%

to 13.5%. While this is the opposite of what we expected, this decrease is not statistically significant.⁶³ In any case, we found no evidence in our data supporting H 1.

3. Likelihood of confusion

The second, third, fourth and fifth hypotheses address aspects that concern the assessment of the likelihood of confusion, namely the impact of the level of attention (H 2) and the level of distinctiveness (H 3) on the rejection rate, the impact of first letter matches on the finding of similarity of trademark signs (H 4) and the impact of the trademark type on the finding of a likelihood of confusion (H 5). In addition to testing these hypotheses, the analysis of the data revealed further findings on the level of attention and the level of distinctiveness. These findings are also presented and discussed below.

a) Level of attention

According to established trademark doctrine and judicial practice, the attention level of the consumer when buying the goods and services for which a trademark is registered is an important factor in the assessment of the likelihood of confusion.⁶⁴ As the general findings show, however, the level of attention is identified in fewer than 40% of all cases.⁶⁵ Yet, the level of attention is mentioned much more often in more recent decisions (see Figure 5), beginning at a rate of below 10% until 2010, rising sharply to 60% of identified levels of attention in 2014 and even peaking further to almost 80% in 2016 and 2018. The reason for the sharp increase starting in 2011 remains unclear, while the second unstable increase is reflected in the IPI guidelines for trademark issues of 2014, which started to provide a series of examples of classes of goods and services for which the level of attention is either high, average or low.

⁶² Sir Ronald Aylmer Fisher's exact test was used to test the independence between alleging non-use and the level of distinctiveness on all cases in which the likelihood of confusion was not significant, Fisher's exact test ($N = 1524$), $p = .13$. For an explanation of this test see: Andy Field, Jeremy Miles and Zoë Field, *Discovering Statistics Using R* (SAGE Publications 2012) 816 and Alan Agresti, *An Introduction to Categorical Data Analysis* (John Wiley & Sons 2007) 45.

⁶³ A chi-squared test was used to test for independence between decision outcome, and the allegation of non-use was not significant, $\chi^2(2, N = 1524) = 1.92, p = .38$. For an explanation of this test see: Field and Miles and Field (n 62) 814, and Agresti (n 62) 34.

⁶⁴ See above, I., IV.2.b, IV.3., V.1.b.

⁶⁵ See above, Table 1, V.1.b.

Table 1. Summary of the assessments of the level of attention, level of distinctiveness and scope of protection. The upper part of the table reports the shares (number) of cases with unidentified and identified level of the respective variable. The lower part of the table reports the shares (number) of the levels in cases in which a level of the respective variable could be identified.

	Level	Attention	Distinctiveness	Level	Scope of Protection
unidentified		60.3% (1345)	31.7% (706)		27% (601)
identified		39.7% (885)	68.3% (1524)		73% (1629)
	low	14.8% (131)	7% (106)	small	3.7% (60)
	average	38.1% (337)	90.2% (1375)	average	93.7% (1526)
	high	47.1% (417)	2.8% (43)	large	2.6% (43)

In addition, there are quite remarkable differences between the levels of attention identified. While the three levels of attention we classified were quite evenly distributed in 2011, the proportion of cases in which an average level of attention was identified rose sharply until 2014 when the level of attention was classified as average in over 35% of the decisions and as high and low in about 12%. Since 2014, however, the share of decisions involving an average level of attention has steadily decreased and the share of decisions involving a high level of attention has sharply increased, while the share of decisions with a low level of attention has remained more or less stable. In 2018, the IPI assumed an average level of attention in about 16%, a high level of attention in more than 40% and a low level in roughly 14% of the decisions.⁶⁶ Again, this sharp decrease is reflected in the IPI guidelines for trademark issues of 2014, in which the IPI mentions several classes of goods and services for which the level of attention is either high, average or low, thereby providing a sound basis for attributing a specific level of attention to the goods and services dealt with in opposition proceedings.⁶⁷

The IPI explicitly states in its guidelines that goods and services of classes 5 (pharmaceuticals), 25 (clothing), 35 (managing and advertising services), 36 (bank, finance, money, real estate, insurance), 42 (scientific and industrial services), and 44 (medical services and medical information services) are perceived with a high level of attention. For goods and services of classes 14 (watches and jewelry), 33 (alcoholic beverages), 34 (tobacco products), and 41 (entertainment and educative services) the IPI assumes an average level of attention. Finally, goods and services of classes 3 (soaps and cosmetics), 29 and 30 (food products), and 32 (beverages) are considered with a low level of attention.⁶⁸

Figure 6 shows that the abstract attribution of a specific level of attention to some classes of goods and services in the IPI guidelines is most often reflected in the IPI's decisions. Indeed, the goods and services from classes which are said to be perceived with a high level of attention (top panel) in the guidelines are overwhelmingly found to be considered with a high level of attention in the decisions. The only exception seems to be class 25.

Table 2. The last row gives the distribution of the levels of distinctiveness found overall ('Total'), and where non-use was not alleged ('no') or alleged but rejected ('yes').

Non-use alleged	Distinct. high	Distinct. average	Distinct. low
no	2.5% (35)	90.6% (1246)	6.9% (95)
yes	5.4% (8)	87.2% (129)	7.4% (11)
Total	2.8% (43)	90.2% (1375)	7.0% (106)

According to the guidelines, however, the goods and services in this class are only considered with a slightly higher level of attention; this subtle difference is mirrored perfectly in the IPI's decisions. Similarly, the decisions concerning goods and services which are considered with an average level of attention (central panel) according to the IPI guidelines are generally attributed an average level of attention in the decisions. Again, there is one exception, namely class 41, which contains services in the fields of education, providing of training, entertainment, sporting and cultural activities. The guidelines do not contain any explanation, but we assume that the divergence in the decisions is due to the broad range of services contained in this class; in particular, we suppose that services in the field of education are usually considered with a higher level of attention than the other services in this class.

Finally, for the goods and services considered with a low level of attention (bottom panel), the general appraisal in the IPI guidelines essentially matches with the attribution in the decisions. Interestingly, however, the IPI diverges quite often from the guideline's assumption of a low level of attention by attributing an average level in the decisions. This is especially true for goods and services in classes 3 and 32. While the divergence might be explained with the broad range of goods and services contained in class 3 (non-medicated cosmetics and toiletry preparations, non-medicated dentifrices, perfumery, essential oils, bleaching preparations and other substances for laundry use, cleaning, polishing, scouring and abrasive preparations), class 32 contains quite homogeneous goods (beers, non-alcoholic beverages, mineral and aerated waters, fruit beverages and fruit juices, syrups and other non-alcoholic preparations for making beverages). At least for the latter class, the reason for the divergence remains unclear. We can only speculate that assuming a low level of attention would lead to accepting a likelihood of confusion in cases in which there is clearly none given

⁶⁶ This finding is in contrast with the results of Fhima and Denvir (n 1) 336, who demonstrated that the EU General Court most often classifies the level of attention of consumers as average.

⁶⁷ IPI guidelines for trademark issues, 1 July 2014, 194 ff; similar: IPI guidelines for trademark issues, 1 January 2017.

⁶⁸ IPI guidelines for trademark issues, 1 January 2019, 264-5.

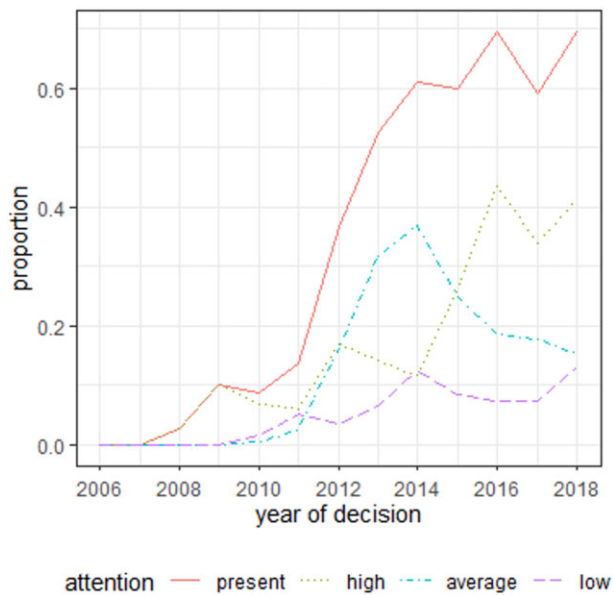


Figure 5. Yearly proportion of decisions in which the IPI identified a level of attention (present) by year of decision. If present, the level of attention is divided into high, average and low.

that consumers actually pay attention to the (often very well-known) brands of all sorts of beverages, including beer and mineral water.

Looking at the empirical findings of the impact of the level of attention attributed by the IPI in its actual decisions on the rejection rate, our data show the expected correlation on which we based our second hypothesis (H 2). While the rejection rate is 23.5% for a high level of attention, it is 18.7% for an average level and only 13.7% for a low level of attention. These numbers are in line with H 2. However, the results are not statistically significant.⁶⁹ For more robust results, future research would have to analyze a larger number of cases. Nevertheless, the mere fact that our analysis has not revealed a pertinent impact of the level of attention on the rejection rate indicates that the relevance of this factor for the finding of a likelihood of confusion is limited.⁷⁰ This should be mirrored in future legal doctrine and judicial practice by moderating the importance of the level of attention for the assessment of a likelihood of confusion.

b) Distinctiveness and scope of protection

Overall, the level of distinctiveness is assessed for 65.5% of all opposing trademarks.⁷¹ Analyzing the yearly findings of some level of distinctiveness over time shows a steady increase from mostly below 50% before 2010 to about 80% in 2018 (see Figure 7). Despite this rising

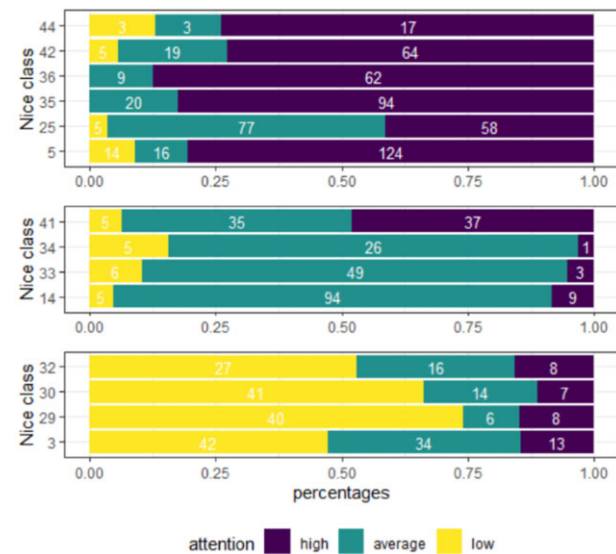


Figure 6. The bars mark the shares of identified high, average and low levels of attention in the IPI's decisions for classes of goods and services which are generally considered with a high (top panel), average (central panel) and low (bottom panel) level of attention according to the IPI guidelines. The numbers indicate the counts in each category. Trademarks for which the level of attention was not identified are not displayed.

number, it is surprising that the level of distinctiveness is not identified for all opposing trademarks given that the IPI guidelines state that the scope of protection of the opposing trademark and, therefore, its level of distinctiveness must be assessed *ex officio*.⁷² The increase over time makes clear, however, that this rule is increasingly implemented and had already reached high levels in 2018.⁷³ With this increase, the finding of an average level of distinctiveness is also rising, whereas the other levels remain roughly stable at low levels.

Strikingly, 90% of all identified levels of distinctiveness are average. According to the IPI guidelines, the IPI conducts its own research in dictionaries, general lexica and on the Internet for assessing the level of distinctiveness.⁷⁴ Our data clearly show that the IPI's base line for assessing the level of distinctiveness is average. Only under special circumstances is a finding of a lower or higher level made. This explains the high number of cases in our sample displaying an average level of distinctiveness.

As trademarks can acquire distinctiveness through intensive and long-lasting use, one might expect that older trademarks would tend to have a higher level of distinctiveness than younger ones. However, the data did not reveal such a correlation.

An analysis of the yearly findings of a distinctive character reveals the following: The rise of the finding of a distinctive character as of 2007 only led to an increasing number of cases in which an average level of distinctiveness was found, while the other levels remained steady with a low share. This suggests that the lack of reference

⁶⁹ A chi-squared test was used to test for independence between decision outcome and level of attention. The test result was not significant, $\chi^2(4, N = 885) = 7.38, p = .12$.

⁷⁰ This finding is in line with that of Fhima and Denvir (n 1) 336, even though they come to an even more pronounced conclusion stating that the analysis of the level of attention of the consumer does not have a significant influence on the finding of a likelihood of confusion.

⁷¹ This number is surprisingly different from the finding of Fhima and Denvir (n 1) 330, reporting that the EU General Court only considered inherent distinctiveness in 32% of the cases and acquired distinctiveness in 18% of the cases, thereby demonstrating a quite remarkable unwillingness to engage in a consideration of distinctiveness and its impact on the likelihood of confusion.

⁷² IPI guidelines for trademark issues, 1 January 2019, 267.

⁷³ Note that identifying a level of distinctiveness is not only determined by the content of the decision but also by our encoding of the latter, which, as mentioned previously, was not reliably coded at a 100% rate.

⁷⁴ IPI guidelines for trademark issues, 1 January 2019, 267.

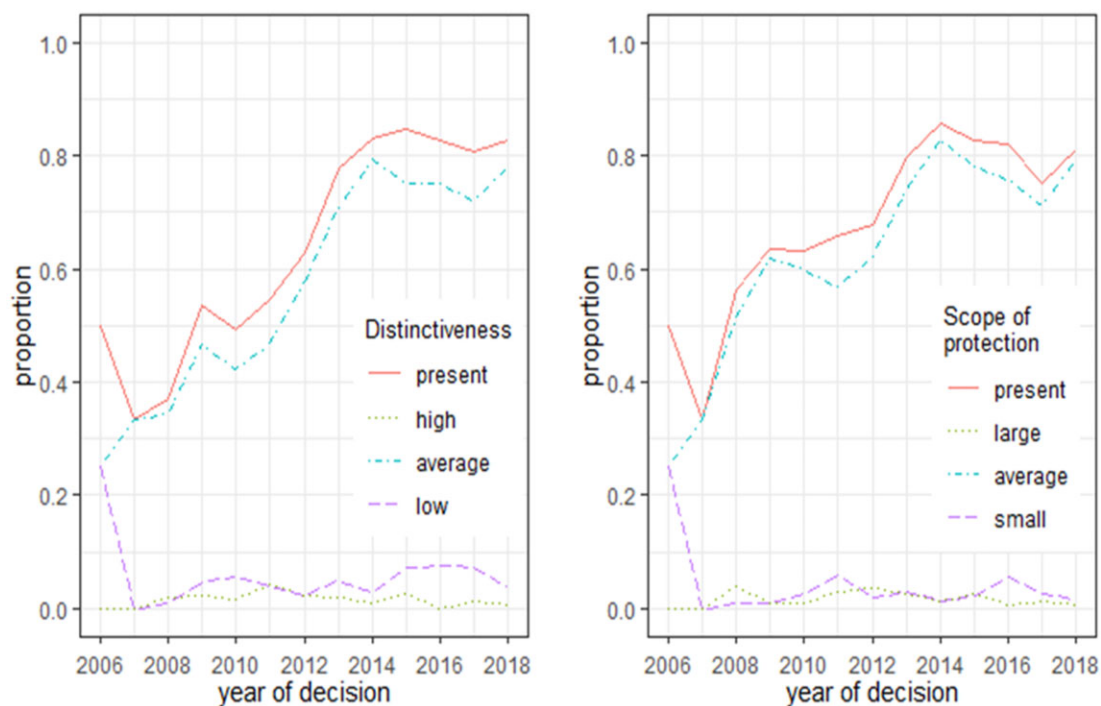


Figure 7. Yearly proportion of decisions in which the IPI identified a level of distinctiveness (left panel) or a scope of protection (right panel) by year of decision. If present the level of distinctiveness or scope of protection was divided in high, average and low, and large, average and small, respectively.

to the issue of distinctiveness prior to 2007 is equivalent to a finding of an average distinctive character.

Our analysis shows that the level of distinctiveness correlates with the outcome of the assessment of the likelihood of confusion.⁷⁵ Curiously, the percentage of oppositions upheld is highest for an average level of distinctiveness (59.9%), followed by the oppositions upheld for a high (44.2%) and a low (15.1%) level of distinctiveness. Thus our data show no evidence for our hypothesis H 3, claiming that a higher level of distinctiveness leads to a lower rejection rate.⁷⁶ While it is in line with legal doctrine that a low level of distinctiveness often leads to a higher rejection rate, it is curious that a high level of distinctiveness less frequently leads to the opposition being upheld than an average one. A potential explanation might be that owners of well-known trademarks may tend to defend their trademarks more aggressively by also initiating opposition proceedings against trademarks that are quite dissimilar. While a case-by-case analysis of the relevant decisions in our dataset supports this explanation, further research is needed to confirm or rebut this assumption.

The number of cases in which the scope of protection is mentioned has steadily increased in the course of recent years, similar to that of the level of distinctiveness (see Figure 7). The scope of protection of the opposing trademark is primarily determined by its distinctiveness.⁷⁷ This is confirmed in the data, with only 18 opposing

trademarks being attributed a scope of protection which is different from these trademarks' level of distinctiveness. A finding of a high level of distinctiveness but an average or even small scope of protection (7 opposing trademarks) is sometimes explained by the IPI by stating that the scope of protection needs to be limited to ensure that the defendant (and others) can use a sign which is in the public domain.⁷⁸ There has been no explanation so far for a finding of a low level of distinctiveness but an average or even large scope of protection (11 opposing trademarks).

It is therefore not surprising to find that – as with the level of distinctiveness – an average scope of protection leads to the highest chance of the opposition being upheld, followed by a large and a small scope of protection, which leads to the lowest chance of success.

c) Similarity of trademark signs

With regard to the assessment of the similarity of trademark signs, we restricted our analysis to the similarity of the typeface of word marks. We tested our fourth hypothesis (H 4), according to which similarity of word marks should be found more often if the first and last letters of the word marks match.

The variables capturing the similarity of the typeface of the word marks are displayed in Table 3.

⁷⁵ A chi-squared test used to test for independence between decision outcome and level of distinctiveness was significant, $\chi^2(4, N = 1524) = 292, p < .001$.

⁷⁶ This finding corresponds with that of Fhima and Denvir (n 1) 330 ff, reporting that distinctiveness has very little impact on the assessment of likelihood of confusion.

⁷⁷ Swiss Federal Supreme Court, BGE 122 III 382, 385; BGer 4C.88/2007, [2008] sic! 44, 46; BGer 4C.258/2004, [2005] sic! 123, 124; IPI guidelines for trademark issues, 1 January 2019, 265; see also Joller (n 7) paras 79 ff, with further references.

⁷⁸ IPI guidelines for trademark issues, 1 January 2019, 264 and the Decision by the IPI in opposition proceedings No. 12461, *CLINIQUE v DERMACLINIQUE BEAUTY FARM*.

Table 3. Occurrences of different types of letter matches in word marks.

Match type	Match	No match	Not available
a... vs. a...	65% (593)	35% (320)	0
ab... vs. ab...	57.5% (525)	42.5% (388)	0
abc.. vs. abc..	50.6% (456)	49.4% (446)	11
abcd. vs. abcd.	39.9% (327)	60.1% (493)	93
...z vs. ...z	52.4% (478)	47.6% (435)	0
...yz vs. ...yz	42.2% (385)	57.8% (528)	0
..xyz vs. ..xyz	35.3% (318)	64.7% (584)	11
.wxyz vs. .wxyz	29.9% (245)	70.1% (575)	93
a..z vs. a..z	28.6% (261)	71.4% (652)	0
ab.yz vs. ab.yz	17.4% (143)	82.6% (677)	93

The case of ‘Vestido’ against ‘Vertigo’, for example, exhibits a letter match regarding the first, the first two and the last letters. In our analysis this counts as a match for the match types ‘first letter match’ (a... vs. a...), ‘first and second letter match’ (ab... vs. ab...), ‘last letter match’ (...z vs. ...z) and ‘first and last letter match’ (a..z vs. a..z), but not for the other types of matches. 593 or 65% of all word mark encounters exhibit a first letter match. A first and second letter match is found in 57.5% of all encounters. These numbers drop to 39.9% for a four-letter match at the beginning of the word marks. The same pattern is found when looking at the last letters of the word marks, but with a lower number of matches, for example 52.4% last letter matches to 29.9% four last letters matches. The first and the last letters match in 28.6% and the first two and last two letters match in 17.4% of all cases. It is surprising that matches in the beginning of words are more frequent than in the end. This might reflect the assumption of the opponents that the beginning of the word is more important for the assessment of the similarity of word marks, motivating them to initiate opposition proceedings.

The success rate increases significantly from 81.7% to 87.4% if the first letters of the word marks are identical. Quite surprisingly however, the success rate does not increase further if not only the first, but also the second, third and fourth letters of the word marks are identical. This finding contradicts the established legal doctrine according to which the beginnings of the words are of particular importance for establishing similarity of word marks,⁷⁹ but it is in line with the findings in perceptual psychology according to which the identity of the first letter is key.⁸⁰ Accordingly, our third hypothesis (H 3), stating that the similarity of word marks is more often established if the first and last letters of the words match, must be made more specific, as the data show that only the matching of the very first letter of the conflicting word marks has a significant impact on the assessment of similarity. These findings correspond with the findings for the similarity of the last letters of word marks. As with the first letters, in cases of identity of the last letters, the success rate increases significantly from 81.7% to 88.1%. And again, the identity of the second, third and

fourth last letters does not lead to an increase of the success rate.

In addition to testing the similarity of word marks, we also analyzed the impact of the trademark type on the results of opposition proceedings. We expected to see relevant differences between different types of trademarks; namely, according to our fifth hypothesis (H 5), we expected word marks to prevail more often than any other type of trademarks.

Indeed, our data show a significant impact of the type of trademark on the result of opposition proceedings. As shown in Figure 8, the success rate is highest if the opponent's and the defendant's trademarks are word marks. This rate decreases quite steadily when moving from the collision of two word marks to the collision of a word mark and a figurative mark containing word elements, to the collision of a figurative mark containing word elements with a word mark and, finally, to the collision of two figurative marks containing word elements.

We assume that the reason for this progression lies in the additional elements taken into account when assessing the similarity between trademarks that contain a word and a figurative element. Indeed, for a figurative mark containing word elements, the IPI needs to find similarities with regard to the word and the figurative element. Interestingly, this finding contradicts the opinion voiced by some, according to which the word element is the most important part in figurative marks containing a word element.⁸¹ Instead, the data are in line with decisions of the Federal Administrative Court and the IPI guidelines, according to which there is no general rule as to the priority of word elements in figurative trademarks containing a word element.⁸²

Interestingly, opposition proceedings between a figurative mark and any other type of trademarks show a much lower success rate than all other proceedings with the exception of opposition proceedings between two figurative

⁸¹ Swiss Federal Supreme Court, BGE 96 II 248; Kamen Troller, *Grundzüge des schweizerischen Immaterialgüterrechts* (Helbing Lichtenhahn 2005) 93; see also Marbach (n 12) para 929 and Christoph Willi, ‘Art. 3’ in Christoph Willi (ed), *MSchG Kommentar, Markenschutzgesetz* (Orell Füssli 2002) para 144, both making reference to (mostly older) judicial practice.

⁸² Decisions of the Swiss Federal Administrative Court, BVGer B-7500/2006 [6.4]; BVGer B-3508/2008 [6]; BVGer B-7367/2010 [6.1.2]; BVGer B-3812/2012 [6.1]; BVGer B-7501/2006 [6]; BVGer B-7768/2015 [5.6]; BVGer B-6173/2018 [3.3]; IPI guidelines for trademark issues, 1 January 2019, 262.

⁷⁹ See above, IV.3.

⁸⁰ Scheier and Lubberger (n 54) 457.

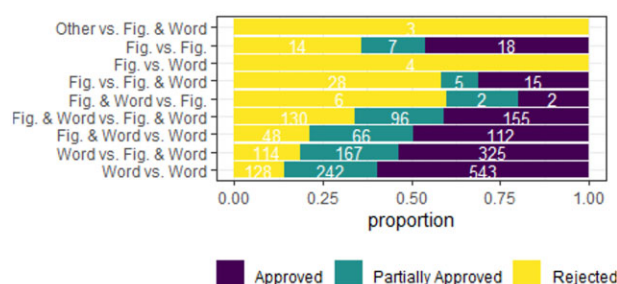


Figure 8. The bars mark the shares of identified rejected, partially upheld and fully upheld decisions for different encounters of trademark types. For example, for encounters featuring word marks against wordmarks 543 encounters were fully upheld, while 242 were partially upheld, and 128 were rejected.

marks, which have a success rate similar to that in proceedings between two figurative marks containing word elements. While the differences for opposition proceedings featuring just one figurative mark (i.e. either as opposing or defending trademark) are quite striking, the number of corresponding cases in our dataset is limited and the results are therefore not robust. Although there may be convincing explanations as to why figurative marks are not similar to other types of trademarks, the low rate of success may also reflect the lack of generally accepted rules for assessing the similarity between these different types of trademarks. Even though first steps to remedy these shortcomings have been taken, further research is needed. In our view, these theoretical deliberations should be based on additional empirical research which should be carried out with larger datasets and thus should be able to provide more robust findings with regard to these relatively rare cases.

More importantly, the sliding scale of the success rates seems to reflect a legal reasoning which is well established in patent law but hardly recognized in trademark law, namely the idea that the scope of protection of an intellectual property right decreases if the number of elements claimed (in patent law, the number of elements in the patent claims) increases.⁸³ Even if this simple reasoning is not reflected in trademark judicial practice and literature, it is clearly recognizable in the facts and should thus be integrated into standard trademark theory.

VI. Conclusion

Our study has revealed a series of interesting empirical results that have normative implications. It also suggests some further empirical research to be carried out on the legal reasoning in the assessment of the likelihood of confusion in opposition proceedings.

Among the findings of our study, four seem particularly relevant: First, our analysis shows that a significant number of oppositions are partially rejected, i.e. that they are only approved for some but not all the goods and services considered. We assume that this is because opponents

often claim a collision with regard to all the goods and services for which the defending trademark is registered. One of the reasons might be the form made available by the IPI. At the very end of this form, the opponents are required to choose between two options: first, ‘The opposition relates to all goods and/or services for which the contested mark is registered’ and, second, ‘The opposition relates only to the following goods and/or services for which the contested mark is registered’. If the opponents choose the second option, they have to specify the classes of goods and services their opposition relates to. Since the choice has no financial consequences for the opponents, we assume that the framing of the questions has a major impact on the opponents’ choice and nudges them towards choosing the first option, even in cases in which there is clearly no overlap for all the goods and services for which the contested trademark is registered. The IPI might want to consider framing these questions in a different way.

Second, our data show that the IPI did not consider the level of attention on a regular basis in older decisions but has done so in more recent ones. Interestingly, the analysis did not reveal a pertinent impact of the assigned level of attention on the finding of a likelihood of confusion. This suggests that the importance of this factor should at least be moderated in legal doctrine and judicial practice. If future research on larger datasets confirms our findings, courts and trademark offices should consider saving the effort of assigning a level of attention. As an alternative, they may want to look for ways to ensure that the attribution of a specific level of attention has a relevant impact on the assessment of the likelihood of confusion.

Third, our research has revealed a surprising difference between another quite well-established legal doctrine and its application in opposition proceedings. According to this doctrine, the beginning and the ending of the words is particularly important for establishing similarity between word marks. Our empirical evidence contradicts this expectation to some extent. In fact, the identity of the first or the last letters of two word marks has a significant effect on the finding of similarity (for first letters: 87.4% vs 81.7%; for last letters: 88.1% vs. 81.7%) but the data show no higher success rates if not only the first and last, but also the second, third and fourth (last) letters of the word marks are identical. This finding calls for a reconsideration of the legal doctrine and/or additional empirical research with regard to the perception of word marks by the average consumer.

Fourth, and last, our data show a sliding scale of the success rates with word marks being the most successful type of trademarks, followed by figurative marks containing a word element, figurative marks and other types of trademarks. As opposition proceedings featuring figurative marks and other types of trademarks are relatively scarce, further research on a larger dataset is needed to gain more robust findings. Nevertheless, the sliding scale of success seems to reflect the fact that the scope of protection of a trademark decreases if the number of elements contained in the trademark increases. Unlike patent law, this simple reasoning is not reflected in trademark judicial practice and literature. As it is clearly recognizable in our data, we suggest integrating this reasoning into standard trademark theory.

While our study is based on a large dataset comprising 2,453 decisions rendered within a period of 16 years,

⁸³ Renate Weisse, *Erfindungen, Patente, Lizenzen* (4th edn, Springer 2014) 23; Marcus Lange, *Anspruchspolitik im Rahmen der Patentanmeldung* (Springer 2006) 122; see also Tobias Bremi, ‘§ 11 Der Schutz von technischen Innovationen’ in Conrad Weinmann, Peter Münch and Jürg Herren (eds), *Schweizer IP-Handbuch* (Helbing Lichtenhahn 2013) para 6.4.

nevertheless, if the aim is for robust results, there are still some limitations. Mainly, we were unable to code the goods and services that are subject to opposition proceedings as the terminology used to describe them is not standardized. As a consequence, we could not assess the impact of this key substantive factor on the outcome of the IPI's decisions. This also rendered the use of a statistical model, e.g. generalized linear models, for the description of the likelihood of confusion assessment impossible. In addition, we had to limit the analysis of the similarity of trademark signs to the similarity of the typeface of word marks. Unlike the similarity of goods and services, analyzing the similarity of other types of trademarks and investigating the impact of a similar sound and meaning

of word marks on decisions taken in opposition proceedings seems possible, although challenging. Both aspects are interesting avenues for future research.

ACKNOWLEDGEMENTS

We would like to thank the Swiss Federal Institute of Intellectual Property (IPI) for providing the data and for the financial support of this project. We would like to thank Zoé Gianocca, Manuela Kälin and Angel Morger for coding, and Alexandra Hansen (MLaw, M.Sc.), Luca Fábíán (MLaw) and Nicole Ritter (MLaw) for research assistance.