



THE SUPREME COURT OF APPEAL OF SOUTH AFRICA
JUDGMENT

Reportable

Case no: 886/2015

In the matter between:

MEDIA 24 BOOKS (PTY) LTD

APPELLANT

and

OXFORD UNIVERSITY PRESS SOUTHERN

AFRICA (PTY) LTD

RESPONDENT

Neutral citation: *Media 24 Books (Pty) Ltd v Oxford University Press SA (Pty) Ltd* (886/2015) [2016] ZASCA 119 (16 September 2016)

Coram: NAVSA, THERON, WALLIS, PETSE and WILLIS JJA

Heard: 30 August 2016

Delivered: 16 September 2016

Summary: Copyright – dictionary – proof of copying – correspondence of example sentences in bilingual dictionary directed at learners aged ten to sixteen – similarity raising suspicion of copying – evidence by compilers of methods used to compile example sentences – lack of access and reference to appellant’s dictionary – explanation for similarity that such inevitable in a work of this type – disputes of fact not resolved by trial – onus of proving copying not discharged on the basis of similarity alone.

ORDER

On appeal from: Western Cape Division of the High Court (Gamble J sitting as court of first instance):

The appeal is dismissed with costs, such costs to include those consequent upon the employment of two counsel.

JUDGMENT

Wallis JA (Navsa, Theron, Petse and Willis JJA concurring)

[1] This is a tale of two dictionaries of a type that many South African schoolchildren, present and past, would recognise. They are bilingual Afrikaans-English/English-Afrikaans dictionaries, the purpose of which is to assist learners using one of English or Afrikaans as a base to obtain a knowledge and understanding of the other language. Both are relatively small with no more than 4 000 to 5 000 entries in each language. These reflect a basic vocabulary incorporating the most commonly used words in each language. Each entry consists of a specific headword¹ and identifies the part of speech (noun, verb, adjective etc), or parts of speech,² of the word; gives the equivalent word in the other language; and provides a short sentence or sentences illustrating its meaning or the different shades of meaning that it possesses. If the entry is in respect of

¹ For example 'dog', 'run' or 'for'. In lexicographical usage the technical expression to describe these words is a lemma.

² For example a word may be both a noun and a verb, as with 'boil' denoting both a type of sore (the noun) and the action of heating some liquid to boiling point (the verb).

an Afrikaans word the sentence illustrating its meaning will be in Afrikaans and then be translated into English and vice versa.

[2] In point of time the first dictionary to be published was published by a predecessor of the appellant, Media24 Books (Pty) Ltd (Media24). It was initially entitled *Tweetalige Aanleerderswoordeboek* and the first edition was produced in 1993. It received favourable reviews as being an advance on what had previously been available. It was republished under the name *Pharos Aanleerderswoordeboek vir Skole* in 2006. For convenience I will refer to it as the *Aanleerderswoordeboek*.³ The respondent, Oxford University Press Southern Africa (Pty) Ltd (OUP) published its work, the *Oxford Afrikaans-Engels/English-Afrikaans Skool Woordeboek* (the *Oxford Woordeboek*) in 2007. It too received favourable reviews. The two are competitors in the market place.

[3] In October 2011 Media24 commenced preparations for a new bilingual Afrikaans-English dictionary to replace the *Aanleerderswoordeboek*. It is customary when a publisher decides to do this for it to investigate the market and examine competing and potentially competing dictionaries in order to decide how to compile its own work in a manner that will be commercially successful in the target market. Although we were not told specifically what first provoked suspicion, Media24's employees came to the conclusion, after looking closely at the *Oxford Woordeboek*, that to a substantial extent it had been copied from the *Aanleerderswoordeboek*. After taking legal advice and

³ Media24 also produces a bilingual dictionary under the name *Pharos Skoolwoordeboek*, which was published in 2005 and is referred to in its application papers. However, all the evidence in support of the complaint of breach of copyright related to the *Pharos Aanleerderswoordeboek*. Accordingly the judgment will refer only to that work.

undertaking further investigations, using independent experts, it launched proceedings in the Western Cape Division of the High Court claiming that OUP was guilty of a breach of its copyright in the *Aanleerderswoordeboek* and sought interdictory and other relief. Gamble J dismissed the application and refused leave to appeal. The appeal is with the leave of this court.

Preparing a dictionary

[4] The preparation of a new dictionary is not a simple matter and involves a substantial amount of planning. Small basic bilingual dictionaries, such as the two in issue in this case, are divided into two sections, one each for words drawn from the two languages. So these two dictionaries have an Afrikaans side and an English side. On each side the first step is to identify the words constituting the core basic vocabulary in that language. These words, which are not necessarily common to both sides,⁴ are the headwords on which each entry is based. The entry translates each word into its equivalent or equivalent in the other language. It identifies the relevant part of speech and frequently used inflections of the word, such as plurals of nouns, or adverbs derived from verbs.

[5] Where a word has more than one meaning all relevant meanings must be identified. If the different meanings arise because the word can be used as more than one part of speech, for example, as both a noun and a verb, they are reflected separately and distinguished by the

⁴ For example the *Oxford Woordeboek* has the English word 'salad', but not its Afrikaans equivalent 'slaai'. Differing views are expressed in the papers on whether this is sound lexicographical practice but nothing turns on it.

identification of the relevant part of speech.⁵ Where the word bears different meanings in different contexts⁶ these must be identified and placed in order, with the most significant and widely used meaning first and the others following in declining order of importance. In a basic bilingual dictionary aimed at learners of one of the languages, only the more significant meanings are given and more subtle linguistic usages are omitted.

[6] Potentially the most significant task in compiling a dictionary is the preparation of sentences or phrases that explain and illustrate the meaning of a word. Not all dictionaries have this. Large-scale explanatory dictionaries of a single language often achieve the same purpose by means of quotations from publications in which the word has been used.⁷ In dictionaries of the type under consideration, the compiler formulates short sentences, referred to as example sentences, illustrating the use of the word in context. Because the dictionaries are aimed at school children it is important that these sentences should be consistent with their life experience so that the meaning may more easily be grasped.⁸ In practice this means that the illustrative sentences will be relatively simple both in structure and theme.

⁵ For example the English word 'desert' may be a noun referring to an arid area of the world, as in the Kalahari Desert, or a verb referring to a person leaving their duty, as in 'the soldier deserts the regiment' or 'the man deserted his wife'.

⁶ For example the word 'recognised' may mean that a person is seen and known to be a particular individual as in 'I recognised my friend John', or may mean allow to speak, as in 'The Speaker of Parliament recognised the Minister of Transport'. Afrikaans examples are the words 'aan', 'af', 'agterkom' and 'almal'.

⁷ The *Webster's Third New International Dictionary* (1985) contains verbal illustrations to illustrate the usage of a word in context and quotations using the word. Similarly the *1913 Webster's Online Dictionary* (available at www.webster-dictionary.org) contains both explanations of the meaning of words and quotations illustrating their usage. Similar examples of this approach can be found in the major dictionaries published by the respondent and its competitors.

⁸ For example there would be little point in illustrating the word 'sang' with the sentence 'The a cappella group sang a Gregorian chant'. A more appropriate sentence would be 'Nkosi sang in the school choir'.

Media24's case

[7] In the founding affidavit of Dr Wanda Smith, herself a lexicographer, Media 24 identified three areas where it claimed that the *Oxford Woordeboek* had copied its *Aanleerderswoordeboek*. They were in the compilation of the headwords or lemmas on both the Afrikaans and the English sides of the dictionary; in the ordering of senses with words having more than one sense; and in the example sentences used in both the Afrikaans and English sections of the dictionary. These latter, it contended, had either been copied directly or had been loosely adapted while following the same themes, sometimes by the alteration of a single word. In support of these contentions it put up an affidavit and report by Dr Anton Prinsloo (to whom I will refer as Dr Prinsloo to distinguish him from Professor Prinsloo who features later in the narrative), an academic from Stellenbosch University, who had examined the entries in the two dictionaries for four letters (B, D, I and S) on both sides and identified where they overlapped in all three aspects. The relevant pages containing those letters extracted from the two dictionaries were attached to Dr Prinsloo's report and highlighted in marker pen. In addition, he compiled two schedules reflecting the example sentences that he regarded as problematic.

[8] Dr Prinsloo's analysis of the areas in which he found commonality between example sentences in the two dictionaries fell into three broad areas. The first instances was where the example sentences were exactly the same. In relation to the word 'baie' both dictionaries included as a sentence illustrating one meaning of the word: 'Baie dankie vir jou help'. The second involved only a small change was reflected as in the case of the word 'brightly' where the two sentences read: 'The sun is shining brightly' and 'The sun is shining brightly today'. Thirdly there were

words where he identified thematic commonality as with ‘behaal’ (‘achieve’) where the sentences were: ‘Christine het baie hard gewerk om sukses in die eksamen te behaal’ and ‘As jy hard werk, kan jy sukses behaal’. I will refer collectively to these different incidents of overlap between the two dictionaries as correspondences. The extent of such correspondences emerged from a statistical analysis undertaken by Professor Martin Kidd, a statistician, the thrust of which was that the degree of overlap between the two works in regard to the example sentences was of the order of 16 to 24% when the statistical analysis was conducted on the basis of a random selection of words.

[9] By the time the matter came before the High Court the complaint in regard to the commonality between the headwords in the dictionaries had been abandoned in the light of the evidence of Professor Daniel Prinsloo, an academic from Pretoria University, (referred to as Professor Prinsloo to distinguish him from Dr Prinsloo, who is mentioned above) concerning the manner in which he compiled the list of headwords used in the *Oxford Woordeboek*. Some argument was addressed to the allegations of similarity in regard to the order of senses in respect of certain words, but little turned on this, presumably because there were few examples of this complaint and a substantial rebuttal of the objection. The focus of the argument, as it was also in this court, was the extent of the alleged copying of the example sentences. That was always the decisive issue. If there had been substantial copying then the claim of breach of copyright had to succeed. If not, it was not suggested that there was sufficient strength in the alleged commonality of sense orders to justify a conclusion of copying of an extent that would warrant the grant of relief.

[10] In making its case in relation to the example sentences Media24 pinned its colours to the mast of the correspondences identified by Dr Prinsloo. It did not rely, as publishers of works such as directories and dictionaries sometimes do, on having set a copyright trap in the work and identifying copying because the *Oxford Woordeboek* fell into the trap. A copyright trap involves the insertion of a false word or false information or a deliberate error in the work, which if it appears in a competing work is indicative of copying having taken place.⁹ Media24's allegation of breach of copyright was based squarely on the correspondences between the example sentences in the two dictionaries. After citing a number of these in his heads counsel summarised the submission by saying that:

'The appellant could continue in this vein indefinitely. The similarities are simply too numerous to consider individually. ... It is submitted that the overwhelming impression that remains is that of copying and the explanation proffered by the respondent of "*common phrases*" and "*saliency*" is simply insufficient to disturb the probabilities in this regard.'¹⁰

So, like counsel, I will focus on the issue of copying in relation to the example sentences. But first it is necessary to deal with the legal framework of the enquiry.

The law

[11] There was no real dispute about the law to be applied in this case. A dictionary is a literary work for the purposes of the Copyright Act 98 of

⁹ As occurred in *Fax Directories (Pty) Ltd v SA Fax Listings CC* 1990 (2) SA 164 (D). There is a famous example of that in the case of the *New Oxford American Dictionary* (2005) which listed the non-existent word 'esquivalence' and gave as its meaning 'the wilful avoidance of one's official responsibilities'. Apparently the word thereafter found its way into a document emanating from the White House.

¹⁰ This echoed his client's view in the founding affidavit that 'Bogenoemde ooreenkomste is bloot té veel en té ooglopend om aan blote toevallige toegeskryf te word'. (The above-mentioned correspondences are simply too many and too eye-catching to be ascribed simply to coincidence.)

1978 (the Act).¹¹ As such the *Aanleerderswoordeboek* was, on publication, a literary work eligible for copyright.¹² The holder of copyright in a literary work is vested with the exclusive right to reproduce the work in any manner or form.¹³ Copyright is infringed where any person, not being the owner of the copyright and without the licence of such owner, does any act that the owner has the exclusive right to do.¹⁴

[12] The application of these provisions was dealt with in this court in *Galago Publishers*,¹⁵ where Corbett JA summarised the legal position in the following terms:

‘Copyright is infringed by a person, not being the owner of the copyright, who, without the licence of such owner, does or causes any other person to do, in the Republic of South Africa, “any act which the owner of the copyright may authorise” (s 23(1) of the Act). To determine the meaning and effect of the words just quoted in relation to a literary work it is necessary to refer to s 6, which provides, *inter alia*, that copyright in a literary work vests the exclusive right to do or to authorize the reproduction of the work in any manner or form, the publication of the work or the making of an adaptation of the work ...

In s 1(2A) it is provided that:

“Any reference in this Act to the doing of any act in relation to any work shall, unless the context otherwise indicates, be construed as a reference also to the doing of any such act in relation to any substantial part of such work.”

Consequently it is not necessary for a plaintiff in infringement proceedings to prove the reproduction of the whole work: it is sufficient if a substantial part of the work has been reproduced. To “reproduce” within the meaning of the Act means to copy and in order for there to have been an infringement of the copyright in an original work it must be shown (i) that there is sufficient objective similarity between the alleged

¹¹ See para (d) of the definition of ‘literary work’ in s 1(1) of the Act.

¹² Section 2(1)(a) of the Act.

¹³ Section 6(a) of the Act.

¹⁴ Section 23(1) of the Act.

¹⁵ *Galago Publishers (Pty) Ltd and Another v Erasmus* [1988] ZASCA 131; 1989 (1) SA 276 (A) at 279G-280C (*Galago Publishers*).

infringing work and the original work, or a substantial part thereof, for the former to be properly described, not necessarily as identical with, but as a reproduction or copy of the latter; and (ii) that the original work was the source from which the alleged infringing work was derived, ie that there is a causal connection between the original work and the alleged infringing work, the question to be asked being: has the defendant copied the plaintiff's work, or is it an independent work of his own? (See *Francis Day & Hunter Ltd and Another v Bron and Another* [1963] Ch 587 at 618, 623, indirectly referred to with approval by this Court in the unreported case of *Topka v Ehrenberg Engineering (Pty) Ltd* 30 May 1983).'

[13] Sometimes the fact that an alleged infringing work has been copied from the original work may be readily identifiable. In the case where it was claimed that the novel *The Da Vinci Code* had breached the copyright of an earlier work, broadly intended to be a work of popular, but possibly serious, historical and religious conjecture, Mummery LJ said:¹⁶

'... it is easier to establish infringement of the copyright in a literary work if the copying is exactly word for word (verbatim or "slavish" copying), or if there are only slight changes in the wording, perhaps in some optimistic attempt to disguise plagiarism.'

That is precisely what occurred in the only reported case I have discovered involving a dictionary.¹⁷ The dictionary in issue in that case was an English-Tamil dictionary, similar in type and purpose to the dictionaries in issue in this case. But the copying relied on was far more blatant than anything that is alleged in the present case. For example it emerged in the course of the trial that the compiler of the infringing work was unable to translate basic English words into Tamil. The conclusion of

¹⁶ *Baigent & another v The Random House Group Ltd* [2007] EWCA Civ 247 para 141 (*Da Vinci Code*).

¹⁷ *Govindan v Gopalakrishna Kone* 1955 AIR 391 (Mad).

the trial court endorsed by the Madras Court of Appeal was that the infringing work was a ‘piratical reproduction’ of the copyright work with:

‘... page after page, word after word, slavishly copied, including the errors and ... the sequence, the meanings the arrangement and everything else practically the same, except for some “deliberate differences” introduced here and there to cover up the piracy’.

It was largely on the same basis that Corbett JA held in *Galago Publishers* that the infringing work was an abridged version of the protected work ‘with extensive language copying’.¹⁸ He concluded that:

‘The similarities are too marked, too many and in too many instances inexplicable except on the basis of copying.’¹⁹

[14] Sometimes, however, copying may be less obvious. As Mummery LJ explained in the same paragraph (para 141) of his judgment in the *Da Vinci Code*:

‘The “text” of a literary work may cover more than the particular words in which it is expressed and extend to its overall content, including the selection, arrangement and development of ideas, theories, information, facts, incidents, characters, narrative and so on.’

The contentions on behalf of Media24 do not involve slavish copying or clumsy adaptive plagiarism. It contended that the copying was more subtle, involving the selection of the subject matter of example sentences and the formulation and arrangement of those sentences. It submitted that the extent of the correspondences was such that it could only have occurred in consequence of repeated reference to their own work.

[15] The reference by Corbett JA to the need for a causal connection between the original work and the alleged infringing work requires a brief

¹⁸ *Galago Publishers* supra at 286D.

¹⁹ *Galago Publishers* supra at 293B-C.

explanation. What the Act prohibits, in the interests of the copyright holder, is the reproduction of a substantial part of the original work. Where it occurs this will usually be deliberate, but that is not a necessary requirement for infringement. The point was made in *Francis Day & Hunter Ltd and another v Bron*,²⁰ referred to by Corbett JA in the passage cited from *Galago Publishers*. In *Bron* it was alleged that Mr Bron had used a well-known tune as the basis for his own composition. He gave evidence that he had never encountered the tune or had cause to play it and this evidence was accepted. But the argument was that as a musician he would, in the ordinary course of events, have encountered the tune and when he came to compose his own retrieved it from his subconscious and incorporated it in his tune. While it was not established that this had occurred, the possibility of unconscious copying was accepted. Thus Upjohn LJ said:²¹

‘There may be cases where, if the circumstances do not justify the conclusion that the defendant, in denying conscious plagiarism, is not telling the truth, they yet justify the conclusion that he must have heard the plaintiff’s tune, and subconsciously reproduced it.’

[16] Any claim that depends to a substantial measure on the notion of unconscious copying must be approached with circumspection. Isolated fragments of memory will not ordinarily suffice because it is only copying of a substantial part of the work that constitutes infringement. As Diplock LJ said in the same case:²²

‘I think that the law is perfectly clear, and that such difficulties as there are in this appeal are solely due to the absence of any factual information about the mental process involved in “subconscious copying”. We know not whether it is rare or

²⁰ *Francis Day & Hunter Ltd & another v Bron (trading as Delmar Publishing Co) & another* [1963] Ch 587 ;[1963] 2 All ER 16 (CA) at 22.

²¹ *Ibid* at 25C.

²² *Ibid* at 27D-E.

common, general or idiosyncratic, nor indeed whether it is possible to remember, not a mere isolated phrase, but a “substantial part of” the remembered work without remembering that one is remembering.’

[17] In the present case, involving the compilation of a dictionary, there can, I think, be no question of unconscious copying. The extent of the alleged copying is necessarily too great to be explicable on that basis. I can understand and accept that, when formulating an example sentence, a compiler might unconsciously make use of a word, phrase or idea that they had recently seen in a similar context after consulting another dictionary. But when that is said to have happened several hundred times it is not explicable on the basis of unconscious copying.²³ Even lexicographers do not bury in their subconscious minds dozens, much less hundreds, of sentences that they have read in other dictionaries, illustrating the meaning of particular words only to pluck them unconsciously from the recesses of memory when they are tasked with explaining the same word. As the authors of a leading textbook in this field note, with a work of any complexity it is unlikely that a person who remembered enough of it to be able to recreate it would be unaware of what they were doing.²⁴ That is especially so in the present case where the copying would have had to occur in the middle of otherwise original work by the compilers. Accordingly, for copying of a substantial part of the *Aanleederswoordeboek* to have occurred in the present case it must have been deliberate copying.

²³ The statistical analysis by Professor Dunne on behalf of OUP showed that the headwords common to both dictionaries numbered between 50% and 60% of the total number of headwords in each. Of these Professor Kidd’s analysis indicated that around 20% contained example sentences that exhibited correspondences according to Dr Prinsloo. The OUP publication had slightly fewer than 10 000 headwords. On these figures therefore between 1000 and 1200 cases of example sentences exhibiting correspondences would be present. For reasons that emerged from Professor Dunne’s evidence that may be a far smaller proportion of the total number of example sentences.

²⁴ Kevin Garnett, Gillian Davies and Gwilym Harbottle *Copinger and Skone James on Copyright* 15 ed (2005) para 7.22, p 380.

[18] Counsel correctly submitted that in the absence of direct evidence of copying – which is usually hard to come by – establishing substantial similarity between the original and the alleged infringing work together with proof of the possibility of access by the alleged infringer to the original work, suffices to raise a *prima facie* case of copying.²⁵ It is then for the alleged infringer to show how its work was produced without copying. A bare denial of having copied is unlikely to displace the inference arising from proof of similarity and access to the original. An explanation of the process adopted in producing the alleged infringing copy that plausibly explains the reasons for the similarity between them will usually be called for. But the onus is not thereby shifted to the alleged infringer. It is always for the claimant to establish copying on the ordinary standard of proof, namely, a balance of probabilities.

OUP's case

[19] Media24's case in its founding affidavit was expressed in broad and general terms. Insofar as the example sentences were concerned the founding affidavit mentioned no specific examples and merely referred to Dr Prinsloo's report and the schedules derived by him from his comparative analysis of the letters B, D, I and S in the two dictionaries. These schedules set out all the correspondences he identified in the Afrikaans and English sections of the dictionaries in respect of those letters. They were attached to the founding affidavit as annexures NB 7 (Afrikaans), containing 215 instances, and NB 8 (English), containing 142 instances, respectively. That involved fewer than 300 words, because in many instances the alleged correspondences related to various

²⁵ *Copinger and Skone James* supra para 7.17.

meanings of the same word. The Afrikaans list referred to 169 out of 712 words common to both dictionaries and the English list to 127 out of 616 common words. The *Oxford Woordeboek* contained 1301 headwords in Afrikaans and 1158 in English in respect of these letters.

[20] The figures in the previous paragraph are taken from the report of Professor Dunne, of whom more anon. Assuming that these four letters were representative of the whole alphabet, so that they can be projected for the entire dictionary, Media24's complaints were confined to example sentences for no more than ten or twelve percent of the words in the *Oxford Woordeboek*. Given that for many words there was more than one example sentence that means that the percentage of example sentences showing correspondences would be markedly less than ten or twelve percent, perhaps nearer five or six percent on Professor Dunne's figures. That was not a great deal but it would probably be sufficient, if there was copying, to say that a substantial part of the *Aanleerderswoordeboek* had been copied. However, it also raised questions about the manner in which that copying had occurred. I will revert to this once I have dealt with the evidence on behalf of OUP.

[21] The answering affidavit on behalf of OUP was deposed to by Ms Megan Hall its publishing manager for dictionaries, school literature in English and general adult books since 2008. She had managed the entire process leading to the compilation and publication of the *Oxford Woordeboek* and therefore was familiar with it in all its aspects. At the outset she noted that it was difficult to ascertain the precise nature and extent of the complaint of infringement from Media24's papers. She also made the point that the dispute, involving as it did a careful comparison between two dictionaries, was not appropriate to be resolved in

application proceedings. OUP's attorneys conveyed as much to Media24's attorneys prior to the commencement of proceedings. Ms Hall indicated that as, notwithstanding this warning, the proceedings had been pursued by way of application any request that it be referred for the hearing of oral evidence would be opposed. As it happened no such application was made and the case was conducted on the footing that it should be decided on the affidavits without resort to evidence.

[22] Ms Hall's affidavit dealt with all the original complaints by Media24 and traversed some matters that later fell away. She explained that Professor Prinsloo from Pretoria University had been commissioned to prepare the lists of headwords, using a corpus derived from OUP's publications. In the light of his affidavit the contention of copying based on the commonality of many of the headwords was not pursued. Three lexicographers, working on a freelance basis, undertook the compilation of the entries in the dictionary. Mrs Aletta Cloete and Dr Liezl Potgieter were responsible for the entries, and hence the example sentences, on the Afrikaans side of the dictionary and Mrs Daphne Paizee prepared all the entries on the English side of the dictionary. At an early stage Dr Phillipus Louw was employed to work freelance as the editor of the dictionary, but after a few months he was employed full-time by OUP as its Senior Editor: Dictionaries. He was responsible for the editing of the dictionary at every stage and where editorial assistants were employed to perform certain distinct tasks he briefed them on the procedures to be followed.

[23] Each of the compilers deposed to an affidavit. Mrs Cloete and Mrs Paizee said that they were sent blocks of about fifty words and were required to compile the entries for those words. This involved

determining the sense divisions of each word and writing example sentences for that word to illustrate its meaning. Both Mrs Cloete and Mrs Paizee said that in undertaking this work they did not have copies of the *Aanleerderswoordeboek* or the *Pharos Tweetalige Skoolwoordeboek*, another of Media24's dictionaries mentioned in the founding affidavit. They said that they did not copy any of their example sentences from either of these works or any other source. They both mentioned other dictionaries that were available to them, none of which were published by Media24. Mrs Cloete explained that her background was as a teacher and an editor of a major Afrikaans dictionary. She had authored an Afrikaans home language textbook directed at grades eight to twelve, which broadly corresponded with the target market for the new dictionary. This background and experience made it easy for her to write simple Afrikaans sentences directed at conveying the meaning of a fairly basic vocabulary to learners aged between ten and sixteen.

[24] Mrs Paizee was equally well qualified to undertake the task of compiling entries for the dictionary. She had been an English teacher in South Africa and the Democratic Republic of Congo, where she had taught English as a second language. She had also been the co-author and author of several dictionaries and held senior positions, including editorial positions, at OUP, culminating in her becoming the Publications Director. She acknowledged the difficulty of pitching example sentences at the right level for the target audience. The only other work she consulted was the *Oxford South African School Dictionary*, a monolingual dictionary aimed broadly at a similar group of learners to the *Oxford Woordeboek*. As mentioned she had sole responsibility for formulating the entries and the example sentences on the English side of the dictionary.

[25] Dr Potgieter was working towards her M Phil in Translation at the time that she assisted with the compilation of the *Oxford Woordeboek*. She therefore had less experience than her co-compilers. Her explanation of how she went about the work was both more extensive and slightly different. As Media24's counsel placed some reliance on this in support of his argument it is best to quote what she said. It appears from the following paragraphs of her affidavit:

'While I did this work as a consultant I had access to and consulted about 12 different dictionaries including the Aanleerderswoordeboek and the Pharos Tweetalige Skoolwoordeboek. I compiled my work using Tshwanelex which is a very useful dictionary compilation programme. Of the 12 Dictionaries I consulted, five were bilingual dictionaries.

The way I worked was that I would first think about the word I was working on in the headword list, its use in language and its sense divisions that would be suitable for the target market of this dictionary. i.e. schoolchildren.

I would then consult the dictionaries I had in my library and check if there were any cases where the words were used in ways different to what I already had in my treatment of the word that might also be necessary to include in a learners' dictionary. I would start by consulting the smaller dictionaries intended for school children and then I would consult the larger ones, like the Verklarende Handwoordeboek van die Afrikaanse Taal (HAT) and other monolingual Afrikaans Dictionaries to conduct the check as set out in the previous paragraphs.

Every word, if it has more than one meaning or sense, has particular meanings/senses which are used more often than others. I would think about which meaning or sense I thought was used most frequently and I would think of example sentences which most typically use that meaning or sense of the word in a way in which a learner would understand. Many words collocate with certain other words (i.e. are frequently used together in phrases or sentences) and I usually included these collocations.

For the purpose of creating my example sentences I would sometimes consult magazines but I would most often use examples from my own life. So for example, if

the word was “tennis”, I thought of my friend who played tennis and I would create a sentence using her name and playing tennis.

I would then translate the Afrikaans example sentence. If I recall correctly I translated all the Afrikaans example sentences I worked on but if I felt unsure doing the translation of any sentence, the Respondent had a translator who did the translation. I believe the English translation I did was checked anyway.

It helped me to translate the Afrikaans example sentences because if the sentence did not work in English then I went back to the Afrikaans sentence to adjust it.

I did not purposefully or consciously copy anything out of another dictionary. It is possible that a sentence or type of sentence may have stuck in my mind but it is not unusual that there are sentences which are similar because it is in the very nature of writing a dictionary that there will be similarities. As I mentioned before many word collocate with certain other words or phrases and are therefore typical combinations that are usually/often used together. These high frequency combinations of words will therefore often be used together in a variety of dictionaries and therefore there will often be similarities in example sentences.

It was very important to me to keep the sentences as short and as simple as possible and to use examples which dealt with something that was relevant to learners’ lives so that they could recognize and identify themselves with the example sentences. I had to be mindful to try to only use words which were already headwords in the dictionary otherwise learners might come across words in the example sentences that they did not understand and were unable to look up in the dictionary.’

[26] In sum, the evidence of the three compilers was that they did not copy from the *Aanleerderswoordeboek*. The possibility that such copying occurred in the editorial process was rejected in the affidavit of Dr Louw. He said that he did not consult either the *Aanleerderswoordeboek* or the *Pharos Tweetalige Skoolwoordeboek* in the process of editing the untranslated work of the three compilers, nor did he make either of these dictionaries available to any of the editorial assistants. Given that the dictionary covered only frequently used words and translations that he knew well, he said he had little need to consult other sources in the course

of his editorial work. When he did so he used a monolingual dictionary, the *Verklarende Handwoordeboek van die Afrikaans Taal*. In the case of subject specific terms in the dictionary, such as geography, he would cross-check translations against an English geography textbook published by OUP.

[27] Apart from the evidence of those directly involved in the compilation and editing of the dictionary, OUP also delivered affidavits by three expert witnesses. The first was Professor Timothy Dunne, an *ad hominem* professor in the Department of Statistics at the University of Cape Town, where he was previously head of department. Media24's expert, Professor Kidd, acknowledged his expertise in the field of statistics. Two points in his evidence were important. The first was an analysis of the correspondences identified by Dr Prinsloo and incorporated in annexures NB7 and NB 8 to the founding affidavit. Professor Dunne obtained the contributor labels from OUP that enabled him to identify which compiler was responsible for each of the correspondences. He analysed that information against the contents of the two annexures and this showed that the correspondences were more or less evenly distributed among the three compilers. Counsel for Media24 attacked this conclusion on the basis that Professor Dunne had not identified the source of his information, but that attack was misplaced as the source was clearly stated in Professor Dunne's affidavit. Media24 did not challenge this evidence in reply. It served to exclude the possibility that the correspondences were all the work of one of the compilers alone – what counsel referred to as a 'rogue compiler'.

[28] The second important aspect of the evidence of Professor Dunne was to draw attention to the statistical principle usually summarised in the

maxim that correlation does not imply causation. What this means is that the mere fact that there is a correlation between two things – in this case the example sentences – does not necessarily mean that the one is the cause or source of the other. In other words, the fact that there is correspondence between different example sentences does not establish that those that came into existence later in time were copied from the earlier ones. That is merely one possibility.

[29] Professor Dunne's analysis pointed to the absence of an identifiable single source of copying (which he referred to as plagiarism). He said that therefore arguably what remained was a requirement that evidence of a collective conspiracy was necessary. His reason was that if the correspondences were to constitute evidence of copying, then either the three compilers colluded to copy from the same sources, while ostensibly working separately, or they all separately, without reference to the others, did so of their own volition.

[30] While Professor Dunne was careful not to venture into lexicographical territory, he drew attention to the need to weigh factors pointing in the direction of copying against other potential explanations of the correspondences. Those explanations were dealt with in the affidavits of Mr Michael Rundell and Professor Taljard, both lexicographers. Professor Taljard identified the constraints that apply to the preparation of a basic bilingual dictionary directed at learners of a language in the context of a school. The headwords in such a dictionary must be carefully chosen by identifying words in frequent use. The profile of the users affects the formulation of example sentences. These must not use language that is more difficult than the word being explained and the examples must relate to the world in which the users

live. As a result, example sentences will tend to revolve around home and school and general matters of common knowledge and experience for learners in the particular age group.

[31] Professor Taljard said that example sentences serve two functions. They not only illustrate word usage, but they also enable the child to decode the meaning of the headword. Thus with the word ‘circle’ the reference is to the fact that a circle is round. The example sentences in relation to a baby will naturally refer to its age so as to identify that it is a baby and distinguish it from other children. In order to illustrate this point Professor Taljard prepared a table in which she compared 59 of the example sentences in the two dictionaries in this case with example sentences from two other similar bilingual dictionaries aimed broadly at the same market sector. The similarity they displayed was apparent from the schedule.

[32] Mr Rundell said that establishing plagiarism or copying in a reference work is more difficult than in the case of a novel or song or a textbook. The reason is that a reference work, like a dictionary, is an assemblage of generally available knowledge and one would expect correspondence between the contents of different reference works dealing with the same topic. The problem is more acute when dealing with a relatively small dictionary aimed at meeting the needs of learners with a relatively low level of language proficiency. Such a dictionary, by its very nature, has a limited range of words and meanings and the necessarily simple illustration of meanings by way of example sentences will call upon well-known common concepts. His conclusion was that: ‘... the resulting texts are likely to show strong similarity: indeed, it would be odd if they did not.’

[33] Mr Rundell illustrated the explanatory function of example sentences²⁶ with a number of examples drawn from Dr Prinsloo's correspondences. He said that:

- The word 'bored' is best explained and understood by referring to someone having nothing to do;
- The word 'broom' is best explained and understood by referring to someone sweeping;
- The word 'second-hand' is best explained and understood by contrasting it with something new;
- The season 'spring' is best explained and understood by saying that it is the season between winter and summer;
- The word 'dear', as a form of address, is best explained and understood by referring to the salutation at the commencement of a letter;
- The word 'drunk' is best explained and understood by reference to the excessive consumption of alcohol.

[34] Like Professor Taljard, Mr Rundell illustrated his point about the likelihood of example sentences of an illustrative nature being similar by way of a short table of 15 words comparing example sentences drawn from five other dictionaries. However, he was careful to say that in his examination of the schedule annexure 'NB 8' there were a few cases where the similarities could not be entirely explained by the constraints on compilers to select the most frequent and typical scenarios. He mentioned five words where he took this view but added that they constituted a minority of no more than 10 instances out of the total of 142

²⁶ This is what Professor Taljard referred to as decoding the meaning of words.

words in ‘NB 8’. He specifically said that he could not agree with the claim by the deponent to the founding affidavit that the similarities are ‘clearly too many and too noticeable to be coincidental’.

Was copying proved?

[35] Media24 established that OUP had access to the *Aanleerderswoordeboek* and that there were sufficiently substantial similarities between the example sentences to raise a prima facie case of copying, calling for an explanation from OUP of how they arose. That explanation was forthcoming in two parts. First, there were explanations by the three compilers of the methods they used in formulating the sentences as well as evidence of their access, or lack of it, to the *Aanleerderswoordeboek*. Second, there was the evidence that such correspondences were likely to occur when creating a dictionary of this type, given its limited range and purpose and the need to adopt an approach to the formulation of example sentences that would fit with the life experience of children.

[36] Media24 chose not to pursue this case by way of trial. Nor did it ask for the matter to be referred to oral evidence. In asking for it to be decided on the affidavits alone, it therefore bound itself to the long established approach described in *Plascon Evans*.²⁷ That meant that the case could not be determined simply on a weighing of the probabilities as they emerged from the affidavits. The facts deposed to by OUP’s witnesses had to be accepted, unless they constituted bald or

²⁷ *Plascon-Evans Paints Ltd v Van Riebeeck Paints (Pty) Ltd* [1984] ZASCA 51; 1984 (3) SA 623 (A) 634D-635D; *Thint (Pty) Ltd v National Director of Public Prosecutions and Others* [2008] ZACC 13; 2009 (1) SA 1 (CC) paras 8-10; *National Director of Public Prosecutions v Zuma* [2009] ZASCA 1; 2009 (2) SA 277 (SCA) para 26

uncreditworthy denials or were palpably implausible, far-fetched or so clearly untenable that they could safely be rejected on the papers. A finding to that effect occurs infrequently because courts are always alive to the potential for evidence and cross-examination to alter its view of the facts and the plausibility of evidence.²⁸

[37] On the issue of copying, Media24 was confronted by the emphatic evidence of Mrs Cloete and Mrs Phaizee that they did not copy at all. At best for it there was an acknowledgement by Dr Potgieter of the possibility that ‘a sentence or type of sentence may have stuck in my mind’ from some other dictionary, not necessarily one of Media24’s range. But a random sentence or two is a far cry from the deliberate copying of which she was effectively, although not expressly, accused. That she denied. Clearly she could not have been a ‘rogue compiler’, as referred to by counsel, because the evidence of Professor Dunne showed that the correspondences were relatively evenly distributed among all three compilers. His evidence reinforced what was already apparent, namely that the correspondences could not be the work of one compiler alone, because they appeared on both the Afrikaans and the English sides of the dictionary. In view of the division of labour among the compilers that necessarily meant that they were the product of more than one person.

[38] Media24’s response to this was a direct challenge to the truthfulness of this evidence. In its replying affidavit Dr Smith said that given the degree of correspondence between the two dictionaries she was

²⁸ *National Scrap Metal (Cape Town) (Pty) Ltd and Another v Murray & Roberts Ltd and Others* [2012] ZASCA 47; 2012 (5) SA 300 (SCA) para 22.

compelled to deny the contents of their affidavits insofar as they alleged that in compiling the example sentences they had not copied from the *Aanleerderswoordeboek*. Its case was therefore that the denials of copying were ‘bald or uncreditworthy denials or were palpably implausible, far-fetched or so clearly untenable that they could safely be rejected on the papers.’ Counsel submitted that this was justified. I do not agree.

[39] It is apposite to mention that counsel was invited to indicate in what manner the alleged copying had occurred. No clear answer to that question emerged in the course of argument. The drumbeat of the argument was simply that the correspondences were so extensive as to rebut any evidence or any probability pointing in the opposite direction. That is not a proper approach. As Harms JA once said:²⁹

‘It is manifestly unfair to argue a case on inferences from some facts and ignoring unchallenged direct evidence to the contrary.’

The circumstances in which a conclusion can be reached on credibility without hearing oral evidence are rare.

[40] The argument for Media24 seems to me to fall into the trap of being misled by what has been referred to as similarity by excision. In *IPC Media*,³⁰ a case that in the way it was conducted bears a close resemblance to this one, the point was made that in order to establish breach of copyright all the evidence must be examined, not only that which points in the direction of copying. Laddie J said:

²⁹ *Dexion Europe Ltd v Universal Storage Systems (Pty) Ltd* [2002] ZASCA 97; 2003 (1) SA 31 (SCA) para 15.

³⁰ *IPC Media Ltd v Highbury-Leisure Publishing Ltd* [2004] EWHC 2985 (Ch); [2005] F.S.R. 20 at para 11 (*IPC Media*).

‘In copyright cases, chipping away and ignoring all the bits which are undoubtedly not copied may result in the creation of an illusion of copying in what is left. This is a particular risk during a trial. Inevitably the court will be invited by the claimant to concentrate on the respects in which his work and the alleged infringement are similar. But with sufficient concentration one may lose sight of the differences. They may be just as important in deciding whether copying has taken place.’

It is important to heed this warning. In order to appreciate this point it is necessary to examine closely the facts in regard to matters where there was no allegation of copying. In other words, in order to deal with the contention that the correspondences alone sufficed to carry the day in favour of copying, we must look at the evidence on behalf of OUP that its dictionary was compiled without copying. Only then can the merits of the argument that we must reject this evidence, based on the correspondences alone, be assessed.

[41] I start with the extent of the alleged copying. Professor Dunne demonstrated that no more (and probably less) than sixty percent of the headwords were common to both dictionaries. Thus the entries in respect of the words that were not common could not have been copied. It was suggested that of the common words only around twenty percent showed correspondences with the example sentences in the *Aanleerderswoordeboek*. These figures showed that the bulk of the work in the *Oxford Woordeboek* was original.³¹ That meant that any copying that occurred was intermittent. An examination of Dr Prinsloo’s schedules does not reveal any discernible pattern and none was suggested to exist, for example, that correspondences occur only in relation to

³¹ Forty percent of the words were not common to both dictionaries and at most twenty percent of the words that were common, that is, one fifth of the remaining sixty percent (twelve percent of the whole), had correspondences in the example sentences. Eighty eight percent of the dictionary must therefore have been original work.

words having multiple complex meanings. Most are in relation to relatively commonplace and quite simple words like ‘box’, ‘broom’, ‘distance’, ‘invite’ and ‘sleep’ to take a few examples. That implied that the alleged copying was both random and intermittent with no apparent common theme, although it was said to be deliberate.

[42] Apart from its reliance on the correspondences Media24 needed to propose a plausible scenario in which the copying it alleged had taken place. As noted in paragraph 39 it did not do so. That complicated its task. If the correspondences, or at least the bulk of them, arose from the three compilers having copied example sentences from the *Aanleerderswoordeboek*, Professor Dunne identified two possible scenarios. The one was that the three had collaborated in copying and agreed to use Media24’s dictionary to that end. However, there was no evidence that they were even aware of the identity of their co-compilers, or knew one another, or were in communication. But that was necessary if the conspiracy theory, as Professor Dunne referred to it, was true.

[43] It was also necessary that these three women, all highly qualified, and two with established records in the field of pedagogical publications aimed at schools and school children, would be willing to risk their reputations and their future careers by engaging in such dishonest actions on a limited scale. They would have had to breach their contracts with OUP in terms of which they warranted that their work was original, and in doing so they would have had to run the risk that a highly skilled and experienced editor might pick up the copying in the course of the editorial process. Given their backgrounds they must all have been acutely aware of the risks involved in copying another publisher’s dictionary.

[44] Finally, what advantage did they stand to gain from doing this? As illustrated by Professor Dunne the correspondences amounted to less than ten percent of the entire task undertaken by them, so random copying would not markedly have eased the burdens imposed by that task. Most of the example sentences were simple and it was unlikely that persons as well qualified as the three compilers would have found it particularly difficult to formulate the sentences. Nor would it have resulted in extra remuneration. It seemed on the surface to be a situation of nothing to gain and everything to lose.

[45] Turning to the second scenario it involved the three compilers deciding separately, and without communication with their colleagues, to copy example sentences from elsewhere and all of them then choosing to do so from the *Aanleerderswoordeboek*. The improbabilities outlined in relation to the conspiracy theory were compounded in the non-conspiracy situation. The reason is that in a conspiracy there may be one person who thinks of engaging in it and then persuades the others to join in. Where there is no conspiracy and they each act separately that requires each of them to have formed separate reasons for engaging in the dishonest and forbidden activity. The probability of one of them doing so was already small for the reasons dealt with in the previous paragraphs. The probability that two of them would do so and would choose the same vehicle to cheat was even smaller. The probability of all three doing so, using the *Aanleerderswoordeboek*, was likely to be vanishingly small.³²

³² If one assumes that each compiler A, B and C was as likely to copy as not to do so, there are eight possible combinations ranging from none of them copying to all of them copying. The likelihood of all of them copying would therefore be 1 in 8. If one of them was less likely to fall prey to the temptation of copying so that given their background and character, the likelihood of that one not copying was 70% and the likelihood of their copying 30%, the likelihood of all three copying falls to 3 in 40. With two such stronger characters it falls to 9 in 200 and if all three are of that ilk it becomes 27 in 1000.

[46] Had Media24 proceeded by way of trial action it would have had to address these questions to the compilers and discredit their evidence through the medium of cross-examination. Instead it argued that the correspondences alone were inexplicable unless copying had occurred. This not only required that the compilers be disbelieved, but it needed to exclude as even possibly credible the explanation for the existence of the correspondences that excluded copying. That lay in the evidence of Professor Taljard and Mr Rundell. They both said that some such correspondences were probable, if not inevitable, when preparing a dictionary of this sort, where there is a rival in the market. The reason was that the dictionaries were basic dictionaries aimed at school children, who were in the process of learning a language. The headwords in the dictionary represented a basic vocabulary and the meanings that the dictionary sought to illustrate were those most common in ordinary usage. The example sentences were therefore directed at being simple and clear and based on the life experience of the children, which was likely to be limited. In the result, so they said, it was not surprising to find correspondences in the example sentences, whether in their formulation or in their context.

[47] There was support for this in the replying affidavit of Dr Smith. She accepted that certain example sentences can illustrate a very typical usage of a word and where there was a correspondence in the two dictionaries in relation to such instances it was not necessarily due to copying. This required a comparison of the sentences and an evaluation of them to see whether their correspondence might be attributable to this. Her only quibble with the argument was that she said that this did not apply to all the example sentences. As regards the evidence of Mr Rundell this was not significantly challenged, although Dr Smith took the

view that his views based on monolingual dictionaries were not entirely applicable to work on a bilingual dictionary. That in turn was based on a lexicographical principle about which there was disagreement, namely whether all the words on the one side of the dictionary should correspond with all the words on the other side. Dr Smith was a firm believer in this, but those responsible for the OUP publication and OUP's expert witnesses did not share this view. Of greater importance was that Dr Smith accepted Mr Rundell's illustrations of how words are used in conjunction with others, and his view that the frequency with which a word is used in conjunction with another is highly relevant to the formulation of a simple sentence to illustrate the meaning and use of the latter were in substance accepted. All that was denied was that this was the reason for the correspondences in the present case.

[48] Similarly much of the evidence of Professor Taljard was accepted and only disputed on minor points of emphasis, nuance and approach, rather than being rebutted on points of substance. In the face of all this evidence directed at showing that copying had not occurred we were invited, from a perusal of Dr Prinsloo's two schedules, to infer that the correspondences were so blatant, and the possibility of their having occurred in the ordinary course of compilation of a dictionary of this type so far-fetched, that we could reject all of this evidence on the papers.

[49] I am unable to reach that conclusion on the basis of Dr Prinsloo's schedules alone and in the absence of the witnesses having given evidence and been cross-examined. Counsel for Media24 accepted that some of his correspondences were inevitable as pointed out by OUP's witnesses. For example, in both dictionaries the sentences in regard to the word 'baba' ('baby') referred to the age of the baby. In both, one of the

sentences in regard to the word ‘blue’ refers to the colour of the sky and another refers to the mixing of blue with another colour to make a third colour. In both, the words ‘alfabet’ and ‘alphabet’ are illustrated by reference to the alphabet having 26 letters. In both, the word ‘afskop’ (kick-off) is explained by reference to the time of the kick-off. In both, the sentence explaining ‘stomach-ache’ refers to eating too much. Whether Mossel Bay or Agulhas is chosen, they are obvious reference points when illustrating the word ‘south’. Describing ‘spring’ as coming between winter and summer is likewise obvious. In each of these cases (and there are many more) it seems to me that the correspondence between the sentences is at least as likely, if not more likely, to have arisen from the adoption of the most obvious ordinary example of the use of the word in common parlance among schoolchildren. In other words, the inference that Media24 asked us to draw from the correspondences on their own was not a permissible one.

[50] If Dr Prinsloo had excluded from his lists instances of the type discussed in the previous paragraph, a clearer picture would have emerged of the cases where copying could more properly be advanced as the explanation for correspondences. That exercise would also have given a better idea of the potential scale of copying, which as already mentioned could not on any basis have been extensive. I could understand evidence showing that in the ordinary course a correspondence of about five percent was to be expected when compiling similar dictionaries, but that correspondence of twenty percent was beyond the norm, but no such evidence was led. Media24 elected to stand or fall by Dr Prinsloo’s comparisons without a trial. In my view it must fall for the simple reason that without the advantages that a trial would have given, it is impermissible to reject the evidence of the compilers and the explanation

that OUP advanced for the existence of the correspondences. That does not involve any final conclusion on the issue of copying. It merely explains why a positive conclusion that copying occurred could not be made on the basis of the correspondences alone.

[51] Two comments remain to be made. The first is that I have not dealt in detail with the possibility that the alleged copying occurred during the editorial phase under Dr Louw. The reason is that the argument centred around the compilers and it was not suggested that, if they were not guilty of copying, Dr Louw and his editorial assistants had discarded some of their work and engaged in copying during the editorial process. The improbability of that occurring was considerable. Not least of many points that could be made is that there would have been no purpose in discarding the work of the compilers for which OUP had paid, in favour of copying from the *Aanleerderswoordeboek*. Again, however, the illumination that cross-examination would have thrown on this issue was absent because of the manner in which Media24 elected to conduct its case.

[52] The second is that the task that Media24 set for itself in trying to discharge the onus of proving copying on the papers without oral evidence was an onerous one. In the three cases referred to in argument where a court rejected a denial of copying by the author of the infringing work³³ that occurred only after a full trial. In *Bron*, *Da Vinci Code* and *IPC Media* the denial of copying was accepted after trial. It seems to me that it will only be in the most egregious case that a court will reject a

³³ *Galago Publishers; Govindan v Gopalakrishna Kone* supra and *Designers Guild Ltd v Russell Williams (Textiles) Ltd* [2001] 1 All ER 700 (HL).

denial of copying merely on the basis of comparing the two works and without cross-examination of the author or authors of the alleged infringing work. It suffices to say that this is not such a case.

Result

[53] In the result the appeal is dismissed with costs, such costs to include those consequent upon the employment of two counsel.

M J D WALLIS
JUDGE OF APPEAL

Appearances

For appellant: A R Sholto-Douglas SC (with him B J Vaughan)

Instructed by: Edward Nathan Sonnenbergs, Cape Town,
Symington & De Kok, Bloemfontein.

For respondent: W R E Duminy SC (with him A E Erasmus)

Instructed by: Cliffe Dekker Hofmeyr, Cape Town,
Claude Reid Inc, Bloemfontein.