

How to find Writers of Vegan Restaurant Descriptions. The Use of Corpora in designing a Language Proficiency Test

Christine Schmailzl

Abstract: In the course of internationalisation, the start-up company “Grünzeug GmbH” wanted to launch their vegan restaurant guide app “vanilla bean” in the United States of America. For the U.S. launch, the app’s database needed to hold descriptions of at least 25,000 different venues. Since the company is growing, suitable personnel for the U.S. launch had to be found. To make hiring more effective, the company wanted to use a language test. This test should assess the applicants’ suitability for writing descriptions of vegan venues in the U.S. A tailor-made test had to be crafted for this purpose. The approach was to use corpora to enhance the validity of language assessment. To design tasks corresponding to the actual job of description writing, a requirement profile was formulated to evaluate the corpus data. This project sets the foundation for future research on the successful combination of corpora and language assessment.

About the author: Christine Schmailzl studied German as a Second Language; Multilingual Consulting; German as Foreign Language and Intercultural Rhetorics at the University of Regensburg; She is currently studying Teacher Training for Secondary Schools at the University of Regensburg, majoring in English and German. The present paper is based on her internship for Multilingual Consulting as well as her studies of English linguistics and was supported by Prof. Dr. Roswitha Fischer.

Keywords: corpora; language assessment; testing; validity; vegan restaurant guide

Corpora can enhance language assessment because they enable test writers to determine linguistic items which are actually used in a domain. The basis of every corpus analysis is a corpus. A corpus is a body of texts which is representative of the domain the linguist wants to analyse. This body of texts is imported into a corpus linguistic comput-

er programme. Such programmes enable linguists to quickly obtain information on the make-up of the texts in the corpus. Corpus analysis, thus, points at what structures are actually used within a given context.

Using corpus-based test content means to assess the proficiency which the test writer attempts to assess. Hence an assessment which is based on corpus analysis is a valid one. The concept of validity is vital to assessment because validity describes that “a test measures what it sets out to measure” (Barker, 2010: 633). Corpora can make language assessment more valid because a corpus analysis shows test writers the most frequent items of real language use within a specific domain. Consequently, data from corpus analysis displays which lexical items, grammatical structures, collocations, colligations and other co-occurrences are actually used in real discourse. For assessing language proficiency, this information is vital – especially if the test is supposed to evaluate language abilities in a specific domain. Having a clear account of language used means having a clear definition of test content. If corpus linguistics can be used for the field of language assessment, it will have a great impact on the validity of language tests, since corpus analysis provides test writers with a clear set of linguistic features. Language tests grounded in corpus data would then not only be more valid than traditionally crafted tests but also fairer to candidates since domain-specific competence can be assessed.

However, corpus linguistic methods have not yet conquered the field of language assessment. The current state of research offers very little on the combination of corpus linguistics and language assessment. Despite the manifold possibilities corpora offer to improve language testing, corpora and the corresponding technologies have not revolutionised the way languages are tested.

This research gap is addressed by this paper: Corpus linguistic methods and considerations about test qualities were combined with a requirement profile for the first time. A domain-specific language assessment for hiring suitable content personnel was designed for the vegan restaurant guide app “vanilla bean”. Assessing general language skills proved obsolete in this case, because specific skills, vocabulary and linguistic patterns are relevant to write descriptions of vegan restaurants. After all, deciding whom to hire based on language assessment can only be effective, if language assessment is tailor-made for the respective purpose. The “vanilla bean” language proficiency test has assessed all skills content workers need.

To grasp the linguistic structures used for vegan venue descriptions in the app, real language usage had to be examined. Turning to corpus linguistics solved this problem. To find out which vocabulary and

structures the applicants needed to know, two different corpora were compiled and analysed, in order to craft the “vanilla bean” language proficiency test: Through both a quantitative analysis of texts displayed in the app and a qualitative analysis of texts which met the standards in terms of length, the test relevant items were determined. Importantly, corpus analysis can only show what to test based on usage frequency. This means that corpus linguistic programmes can scan the respective corpus for the most frequent nouns, verbs, collocations, word combinations and the like. After performing the corpus analysis, the test writer arrives at lists of frequent items.

What corpus data does not show the test writer, however, is *how* to test these structures. In other words: Corpus analysis can only provide data and not the way of using this data for language assessment. To decide how to test the relevant linguistic structures, a requirement profile was used to evaluate the corpus data. A requirement profile is a systematic overview of all skills and subskills needed to perform a task. In the requirement profile for the “vanilla bean” language proficiency test, the process of writing restaurant descriptions was operationalised in terms of content. The clear account of the requirements acted as the basis for evaluating the corpus data. With the help of this requirement profile, the tasks of the test were designed in a manner that corresponds to the job of a “vanilla bean” content worker.

The context of the research project

“Grünzeug GmbH”, founded in Regensburg in 2015, aims to make eating environmentally friendly easy for everybody. Not every restaurant today offers vegan options. Therefore, it can be hard to follow an eco-friendly lifestyle on the go. This is why the start-up company developed the smartphone app “vanilla bean”, a restaurant guide which only lists venues with at least one vegan dish on the menu. The app provides its users with the service of finding vegan food options close to them or in a city of their choice. In addition, users can search for venues which offer gluten-free, organic, fair-trade, lactose-free, raw, vegetarian and omnivorous options. Moreover, the app provides its users with further information on the venues, such as address, opening hours, Facebook page and website, telephone number, price range, e-mail address, photos of the venue, and the food served as well as a slogan and a description of the restaurant.

After the launch in Germany on the 9th of December 2015 and in Switzerland on the 15th of March 2016, the launch of the app in the first

English-speaking countries followed in 2016. On the 15th of September 2016, the app's service was made available to people across England, Ireland, Northern Ireland, Scotland and Wales. By virtue of being a growing start-up, "vanilla bean" has to constantly extend their user base. Thus, entering new markets is crucial. At present, the United States of America is the biggest English-speaking vegan market: Plant-based alternatives have experienced an increase by over 8 per cent in the past year alone – with a rising trend.¹ Consequently, the next step to make the app an international success was to establish it in the United States on the 23rd of January 2018. The project presented in this article concerns the U.S. launch.

The problematic point for the U.S. launch was the acquisition of content employees. As the app provides data not only on the listed venues but also on restaurant descriptions, the entries of the respective venues cannot be generated automatically. To successfully introduce the app to the U.S. market, the database needed to hold descriptions of at least 25,000 venues. To make recruitment as efficient as possible, "Grünzeug GmbH" wanted to use a language assessment test to decide whom to hire for the U.S. launch. Consequently, the big question was the following: What does the perfect test for this purpose look like?

The test qualities of objectivity and reliability

Faced with the task creating such a language proficiency test, two more specific questions arose. The first concerned the test content: What needs to be tested to allow conclusions about the applicant's suitability for writing descriptions? The second concerned the actual make-up of the test: How can such a language test be crafted? A domain-specific test has to be designed on the basis of a well-grounded theoretical framework.

Importantly, assessment informed by corpora "should [also] be subject to the normal requirements of validity and reliability" (Alderson, 1996: 258). This again fosters the proposed link between theories of language assessment and corpus linguistics. Corpus data on its own does not enable test writers to create useful language tests. Thus, it was crucial to see what makes a test a *good* test. According to test qualities, every test has to be objective, valid and reliable. A test's objectivity describes whether its results are independent of the respective rater (Lienert / Raatz, 1998: 7). Hence, objectivity clarifies two things: First-

1 "Plant Based Foods Sales Experience 8.1 Percent Growth Over Past Year", in: <http://www.prweb.com/releases/2017/09/prweb14683840.htm>.

ly, the respective person who conducts and rates the test does not impact the outcome. Secondly, the test is the same for all test takers, meaning that each test taker has the same amount of time to complete exactly the same tasks with the same means. Thus, it does not play a role whether rater A or rater B conducts, corrects and/or marks the test: The results will be the same, because all participants take the same test under the same conditions.

Ensuring objectivity was the first step in creating the “vanilla bean” language proficiency test, because objectivity is a prerequisite for all other test qualities. If a test is not objective, it will be neither reliable nor valid. The best way to achieve high levels of objectivity of application is to minimise the interaction between the participants and the observers (Lienert / Raatz, 1998: 7). To eliminate observer influences, an online test was used. Since the test takers do not get instructions by a human observer but in written form only, there is no way the test procedure could be influenced: Every test taker answers the test’s questions on their computer in the same online form. Only closed tasks were used, because they are fully objective, as specific answers are expected from the test taker. Each item was either correct or incorrect. Every correct answer earned one point. In case one answer included two aspects, this answer was counted twice. In the case of an online assessment, the process of correcting and scoring can be computerised. Automated rating of closed tasks not only saves time, because it does not consume human resources, but it is also error-free. To ensure the same test conditions, the tools used to take the test, the test itself, and the time to answer the assessment had to be the same for each applicant. An adequate time frame to answer the respective tasks was determined in the course of evaluating the pre-test. The pre-test takers were asked to note down the time which they needed to complete each task. The average time that the pre-test takers needed to answer the assessment was used for the final test.

Concluding, the “vanilla bean” language proficiency test is objective because all test takers have the same time frame to complete the online assessment comprised of the same tasks. All applicants take the test under the same conditions. There is no interaction between the observer and the test taker because all test content, i.e. both instructions and tasks, is provided in written form only.

Having ensured objectivity, reliability has to be achieved. If a test is reliable, its “results are accurate, consistent and dependable” (Barker, 2010: 633). A reliable test provides the observer with error-free results that are persistent over time. To ensure reliability, the method of comparing two parallel tests was used. This was possible, because the “vanil-

la bean” language proficiency test includes two C-tests. C-tests are a special kind of gapped task format in which the last letter of every second word is deleted. The person taking the test has to fill in the gaps. C-tests ask test takers to use their communicative as well as their linguistic knowledge of a language. Therefore, C-tests help to draw conclusions about general language proficiency (Karimi, 2011: 7). This task format tests productive knowledge of content words, lexis, inflection, orthography as well as overall text comprehension and coherence.

The C-tests in the “vanilla bean” language proficiency test were supposed to assess domain-specific language. Thus, attention was paid to deleting letters of key words, i.e. words that are important to the field of vegan restaurant descriptions. Both C-tests are of comparable difficulty and length. The performance of the pre-test takers in these C-tests was compared to ensure inner consistency. The pre-test takers achieved similar scores in both C-tests, although one C-test was set at the beginning of the test and one at the end. The order of the C-test, then, did not influence the test result, proving that the “vanilla bean” language proficiency test avoids falsification of the results due to task order. Having excluded possible sequence effects, the results of both C-tests can be used to examine inner consistency. The performance data of the test takers in both C-tests show high correlation. Hence, the language test has inner consistency, which makes it reliable.

Validity and corpus linguistics

The most problematic test quality is validity. A test’s validity is concerned with whether the test measures the content that it intends to measure (Akbari, 2012: 30). Validity can only be assessed content-wise: Is the test fit for its purpose? The relationship between test content and test purpose is crucial. A test which is perfect for one domain can be utterly useless for another. Depending on the purpose and domain, the linguistic structures used differ dramatically. Thus, a test can only be valid, if its conception is tailor-made for its field of application. If a test is supposed to tell the observer whether the test taker’s proficiency of English is high enough to work in a certain position, a valid test will tell the employer who to employ for that job. However, decisions based on assessment can only be justified, if the test score really reflects the test taker’s proficiency and not anything else (He / Dai, 2006: 377). In fact, validity describes whether a test is a suitable basis for decision making (Messick, 1989: 13).

The “vanilla bean” language proficiency test was supposed to determine whether an applicant will be a good content worker. The test had to capture one’s aptitude for writing restaurant descriptions. Using a test that attempts to measure general language proficiency in English, would therefore not allow drawing conclusions about an applicant’s aptitude to work for the app. Evidently, a test which perfectly fits its purpose, i.e. finding good restaurant description writers for the U.S. market, had to be designed. The difficulty was to define the domain-specific language skills that are needed for writing descriptions of vegan venues. Which items need to be included in the test to find suitable employees?

Corpus analysis can only provide test writers with useful data, if the corpus design matches the intended purpose of the test (Alderson, 1996: 254). It is crucial to choose a suitable compilation of texts for the data to be representative of the respective language domain. Depending on the domain of language use, the corpus’ make-up has to be different. If, for example, one is to assess the test takers’ capability of business English, the texts have to be from that domain. Domain specific corpora bear great potential for determination of test content for specific purposes (Barker, 2006: 3). The use of domain-specific corpora, however, has not yet become widespread, even though they are a pre-requisite for the development of “valid and reliable domain-related tests” (Taylor et al., 2003: 250).

Using corpora to generate test content for the “vanilla bean” language proficiency test

Before writing a test, the test writer has to clearly define the content of the assessment (O’Sullivan, 2012: 82). The basic assumption was that the data gained via corpus analysis of the existing texts would ensure the validity of the language test. Two corpus analyses of two different corpora were conducted. Through both a quantitative analysis of texts displayed in the app and a qualitative analysis of texts which met the standards in terms of length, it could be determined which linguistic items to test.

Since the app had already been launched in the United Kingdom and Ireland, a large number of restaurant descriptions in British English was available. The fact that the texts were written in British English was not a problem, because orthography and lexis can easily be adapted to American English. At the time of creating the test, the database held 973 discrete English texts, which I used for the first corpus analysis.

The most frequent words of this large corpus were found via the concordance programme “AntConc”. “AntConc” is computer software which enables the user to scan texts for key words in context and to analyse them.² “TagAnt” is an additional programme and can be used together with “AntConc”. It provides information about word classes for each linguistic item in the corpus. The text files tagged with “TagAnt” were imported into “AntConc”. Lists of the most frequent words according to word classes were created. The word list provided a clear amount of the vocabulary for the descriptions of the app. With the Pivot table function of the spreadsheet programme “Excel”, the individual lists of word classes were then ranked according to their frequency. The categorization into word classes was important for two reasons: First, the individual word class lists provided an overview of the distribution of the vocabulary used for restaurant descriptions in “vanilla bean”. Secondly, these lists yielded information about the grammatical structures used in venue descriptions. This quantitative analysis of the existing restaurant descriptions helped to make the first, vital decision in the test crafting process: There are frequently used phrases in the venue descriptions. Hence, it would not make sense to test vocabulary and grammar separately. All tasks have to be constructed following lexico-grammatical principles. Lexico-grammar recognises syntax and lexis as interconnected parts of speech (Sinclair, 1991: 104). Knowing these expressions or chunks of words will make the production of venue descriptions easy, because they function as flexible components to compose correct utterances.

The second corpus analysis provided a qualitative overview of the make-up of the restaurant descriptions. For this corpus, only texts with the length of 350 to 450 characters were used. This number of characters was chosen, because “Grünzeug GmbH” considered it as optimal. The second corpus is significantly smaller than the first one. It comprises only 106 different texts. However, these texts are of the expected quality in terms of length and style. To analyse the second corpus, the data analysis programme “MAXQDA” was used. “MAXQDA” is a computer programme used for both qualitative and quantitative analyses of texts.³ To gather the crucial data, the feature “MAXDictio” was used in order to determine the most frequent word combinations. These combinations were then integrated into the language proficiency test as items.

² “AntConc,” Laurence Anthony, 2018, <http://www.laurenceanthony.net/software>.

³ “MAXQDA Plus (incl. MAXDictio),” MAXQDA, 2018, <https://maxqda.com/products/maxqda-plus>.

Importantly, corpus analysis can only show what to test in terms of usage frequency. Corpus analysis does not generate ready-to-use assessments. How can language assessments be crafted with corpus data? First of all, it is important to note that the content workers need to communicate in a specific domain, namely, the domain of food and drinks with a focus on vegan, gluten-free, organic and regional offers. Secondly, not only the domain-specific language has to be assessed. The writers for the app need to be able to perform all the tasks related to the production of a description. The tasks in the test have to correspond to the requirements of the job. Being aware of the multiple aspects that need to be included, the question was how to define and operationalise these aspects.

Making use of corpus data with the aid of a requirement profile

To create appropriate test items, the data from both corpus analyses was evaluated with the aid of a requirement profile. The use of a requirement profile ensured the precise definition of the test construct. The tasks were corpus-based and designed in a way that assessed the test taker's suitability for the job of a writer of restaurant descriptions.

The requirement profile encompassed all skills needed to write descriptions. In general, every content worker needs to adhere to the company standards for restaurant descriptions. The "vanilla bean" guidelines are made available to the subjects before taking the test. It is expected that the applicants grapple with these guidelines. Another basic requirement is the ability to do research. To write the venue descriptions, the content workers need to find information on the respective venue online. It is often necessary to use various resources, such as the venue's website, its menu or its Facebook page. Importantly, the available information needs to be evaluated in terms of relevance. Lastly, all the pieces of information need to be put into a coherent text: The description should provide the user with a good idea of a restaurant's vegan, dietary and, if any, its sustainability concept. The approximate text length should be 4 to 6 sentences.

The organisation of the description has to adhere to the "vanilla bean" guidelines. Above the running text, the app user has to be notified of possible obstacles that could make the restaurant visit a let-down under the "TAKE NOTE" caption. To do this, content workers need to be able to perceive such possible obstacles, i.e. varying opening hours or kitchen times. All essential information on the venue has to

be included in the introduction sentence. This sentence encompasses the name of the restaurant, its location, the type of cuisine, i.e. Italian, fast food, clean eating or raw, the type of eatery, i.e. café, restaurant, ice cream parlour, etc., as well as its diet, i.e. vegan, vegetarian, vegan-vegetarian or omnivorous with vegan options. Following the introduction, the writers have to give examples of vegan dishes on the menu – provided that the venue’s menu is accessible online. It has to be highlighted if the venue offers more than just regular food, i.e. mock meats, raw vegan or gluten-free dishes. Further information on the venue can be included afterwards. For example, the text can give information on whether sustainability is an important aspect of the venue’s concept: Does the restaurant use organic, fair-trade or regional ingredients? In case the venue regularly hosts events, i.e. Sunday brunch or live music evenings, this should be mentioned at the end of the description.

In addition to fulfilling the content-wise requirements, the applicants have to have all the markings of the needed language ability. In terms of grammatical items, they have to be capable of using correct punctuation, syntax, 3rd person inflection, collocations, simple present as well as American English orthography and lexis. Moreover, they have to be able to use paraphrasing strategies, such as finding synonyms, splitting one sentence into two, relative clauses, converting sentences from active to passive and vice versa, nominalisations, verbalisations and substituting nouns with pronouns and vice versa.

Pragmatics plays an important role, too. The crucial aspects here are a neutral style and relevance. To present the information on the respective venue as neutral as possible, exaggerations, comparatives and superlatives need to be avoided. The descriptions have to be written in the 3rd person or passive. The personal pronouns “I” and “we” are not to be used. Importantly, the information given has to address the app’s target groups. The descriptions only give information on vegan options. It is not interesting for “vanilla bean” users to know about the restaurants non-vegan dishes or ingredients, for example.

Tasks design of the “vanilla bean” language proficiency test

In the “vanilla bean” language proficiency test, the tasks were conceptualised to assess the individual skills needed to perform the job of a writer of vegan restaurant descriptions. The requirement profile was not only used to evaluate the corpus data but also to choose suitable task formats. Thus, the requirement profile helped utilizing the word frequency lists and lists of word combinations for test creation.

Each task focuses on one specific skill to ensure that exactly this particular skill is measured. Gap fill activities ask the test taker to insert the correct word or phrase either on their own or choosing from a selection of options. In the “vanilla bean” language proficiency test, gap fill activities were used to assess linguistic correctness, namely syntax and grammar, as well as content and order of description texts. Pieces of information had to be sorted into a grid. In another gap fill activity, sentences in the passive voice had to be put into the active voice and vice versa. Since stylistic variation plays an important role when writing descriptions, there was a task on choosing the correct synonym for the underlined part of the sentence.

Multiple-matching tasks demand to order items into the correct sequence or to match corresponding items (Barker, 2010: 635). In the “vanilla bean” language proficiency test, multiple-matching tasks were used to test knowledge on collocations, research skills, morphology, pragmatics, different cuisines, and to check whether the test takers are able to comply with the company’s set order of description content. To assess the latter, the test takers had to put pieces of a restaurant description into the correct, chronological order (Task example 1).

The text used for this task is a model description. Each sentence includes a different aspect of venue descriptions displayed in the app. However, the order of the sentences is incorrect. The applicants had to sort the given sentences content-wise. Thus, the task assessed whether the applicants had familiarised themselves with the company’s guidelines and whether the applicants were able to apply these standards to written texts. Another multiple-matching task requested the test takers to sort pieces of information into categories with regards to relevance (Task example 2).

3. It's important that the description of each venue follows a logical order. Place the following parts of the text in the right order.

Drag items from the left-hand list into the right-hand list to order them.

There are vegan and gluten-free options, such as stuffed bell peppers, sweet potato skins, falafel wraps and fruit salads. →

The View Deli is called like this for a reason: →

Not only hot drinks and soft drinks but also a variety of different dishes are served. →

Since the café is located on level 4 of the Students' Union building, you can enjoy the great view over Sheffield. →

The View Deli is a meat-free café, which offers Indian-inspired dishes. →

Task example 1: Structuring venue descriptions correctly

24. Not everything you find on the websites of venues is important. Thus, there is no need to include every piece of information into the description of a venue.

Tick the correct boxes.

	A must	Nice to have	Not relevant
Mary Ward Center Cafe is a vegetarian cafe, which serves Western and American cuisine.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The café functions as art gallery.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The café only serves free-range eggs from local farmers.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
All food is also to take-away and the café offers catering services.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The meat served is sourced from producers that ensure high standards of animal welfare.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
There are vegan options and they are labeled as such.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Task example 2: Distinguishing important from unimportant pieces of information

In this task, the applicants were asked to decide whether the given pieces of information matter to the services the app provides. As the users of “vanilla bean” are only interested in information concerning vegan food and ingredients, the applicants had to categorise information on non-vegan food and ingredients as “Not relevant”. Furthermore, the test takers had to decide whether the respective piece of information has to be included in every venue description or whether the respective piece of information only adds more value to a venue’s de-

scription. Pieces of information which are not compulsory had to be categorised as “Nice to have”.

In addition to multiple-matching tasks, multiple-choice tasks were included in the assessment. Multiple-choice tasks provide a set of possible answers where more than one answer can be correct. (Barker, 2010: 635). The “vanilla bean” language proficiency test included a multiple-choice task to assess whether the applicant can read menus in terms of gathering information which is relevant for the target group. On the basis of a menu, the test takers had to decide whether the given statements were true or false. Another multiple-choice task assessed collocations, content words, compounds and prepositional phrases. Furthermore, there was a task in which the test taker had to decide whether the given sentence is appropriate for the app or not.

Error spotting and correction tasks test both receptive and productive skills. In the “vanilla bean” language proficiency test, linguistic correctness, i.e. orthography, inflection, irregular verbs, American spelling and punctuation, was assessed with an error spotting task. The task consisted of a venue description with the most frequent words and linguistic structures determined by the two corpus analysis. Mistakes were then added to this text. Besides linguistic correctness, this task format also tested how well the applicants can proof-read and how meticulously they work.

To ensure that the tasks were valid and the instructions were clear as well as to determine the points and time to pass the test, a pre-test was conducted. The paper pre-test was comprised of 12 different tasks. 18 people took the pre-test, 6 of them were company employees and the rest university students. The pre-test takers had to solve the tasks and stop the time they needed to complete each task. After each task, they had to fill out a feedback form on the task’s clarity, length and level of difficulty. After completing the whole test, the pre-test takers had to fill out a feedback form on their overall impression of the test: Was the level of difficulty appropriate? Was the length appropriate? Moreover, the pre-test takers could make suggestions on how to improve the language test.

The performance data was gathered for each test form, each task and each item within the tasks. Tasks which did not deliver results that could be used to determine the test taker’s suitability as a content worker were removed. This means that tasks which were answered falsely or correctly by all test takers did not make it into the final test. In addition, the instructions were made clearer with the aid of the test takers’ feedback. The number of points to pass the assessment was set by looking at the pre-test takers results. Since the test writer cannot

fully estimate the level of difficulty beforehand, the pre-test performance data allows conclusions to be drawn. Based on the pre-test data the level to pass at was set at 70% of the points.

Since the “vanilla bean” language proficiency test was supposed to be an online test, a way to transfer it from paper-based to a digital format had to be found. There are a number of free-to-use online survey tools. After evaluating the options, the company decided to use “Survey Gizmo”. “Survey Gizmo” was the only tool which fulfilled the needs: It offers ready-to-use task formats and a timer. Moreover, the tool is easy to use, not overpriced and has an appealing graphical user interface. Transforming the “vanilla bean” language proficiency test from paper to digital went quickly after familiarizing oneself with the tool. None of the applicants had problems taking the test.

Corpora and language assessment: A successful union?

The research project started with the question of how to craft a language assessment to hire personnel for the U.S. content department of “vanilla bean”. The crucial test quality to ensure was validity. A valid recruitment test for “vanilla bean” would predict the applicants’ suitability to write restaurant descriptions. The hypothesis was that corpus-based test content would guarantee the validity of the assessment. To combine the fields of corpus linguistics and language assessment, a requirement profile was used. This requirement profile shaped both the evaluation of the corpus data and the task design. A domain-specific language test for professional purposes was developed. Could the hypothesis that corpus linguistics and language assessment is a successful union be proven?

In July 2017, “Grünzeug GmbH” used the language assessment for the first time. According to the U.S. launch manager, Dr. Jenny Di Leo, the test was crucial in finding suitable employees since the test made the process of acquiring writers both objective and efficient. After minimal advertising of the jobs, 20 people applied as U.S. content workers. Based on the test scores, eight of them were hired. Seven of the new employees were able to produce good content from the start. Only one person had difficulties at first. These difficulties, however, could be managed with extra coaching. After their probation period, all content workers kept their job and were offered more hours. Two content workers went onto a permanent role after completion of their fixed-term contracts. Hence, the language test’s validity was proven by the success rate in recruitment.

The “vanilla bean” language proficiency test enabled the company to find the best candidates for the job rather than having to solely rely on application documents: Candidates with a less appealing curriculum vitae achieved high scores on the test. These have proven to be competent and reliable employees. In addition, the language test helped to save a lot of time, which could be used to build up more restaurant data than anticipated.

In October 2017, the “vanilla bean” language proficiency test was used again as “Grünzeug GmbH” wanted to hire more content workers. Again, it proved to be a helpful tool. New employees could be acquired in an efficient way. Evidently, the language test allows accurate conclusions in terms of the test takers’ aptitude to create content for the app “vanilla bean”. Today, the language proficiency test is still used to acquire new employees. At the same time, the test functions as a first training session for future employees. Since answering the tasks demands good knowledge of the company’s standards for restaurant entries, the applicants have to closely read the guidelines. Therefore, it takes less time to introduce new employees to the use of the app’s database.

The project’s findings clearly indicate that the starting hypothesis is true. It makes sense to use corpus linguistic methods to design assessments. The use of corpora in combination with a requirement profile enables test writers to design domain-specific language assessments. Basing the “vanilla bean” language proficiency test on corpus analysis made it possible to craft a valid recruitment test.

Corpus-based assessments – the future of language testing?

Having established that the approach used to craft the “vanilla bean” language proficiency test worked, the question whether this approach can be applied to other projects arises. Starting with a corpus analysis of a domain-specific corpus, the data would be evaluated with a requirement profile. The assessment itself would have to be subject to considerations about test qualities. These steps are not restricted to the field of application presented in this paper. This approach to designing language assessment can be considered universal. Hence, it will be possible to transfer the approach to other domains and purposes. The applications of corpus-based language tests are countless: They can be used in small, medium-sized and big companies which want to hire employees in specific departments. Moreover, corpus-based language assessments can be used within various scientific domains. A university

could use such a test to evaluate the applicants' competence in academic English within the branch of linguistics, for example. In addition to recruitment processes, corpus-based language assessments could be used in the context of the Common European Framework of Reference for Languages. The can-do descriptions given for the respective levels of proficiency could be supplemented with concrete linguistic structures. Corpus analysis could make the CEFRL more transparent because the data on vocabulary and grammar for different levels of proficiency could be given.

In conclusion, corpora could revolutionise how languages are tested. Basing test items on real language use enhances the validity of the respective assessment. Programmes for corpus analysis provide the tools to quickly obtain crucial data on the make-up of domain-specific texts. Importantly, test writers have to find a way to utilize this data to design language assessments. The project presented in this paper can be seen as a fundamental first step: The gap between corpus data and test items was bridged by a requirement profile. Future research should evaluate the proposed approach. Applying this approach to other projects attempting to assess domain-specific language will help to investigate its efficiency.

References

- Akbari, Ramin (2012): "Validity in language testing", in: Coombe, Christine et al. (Ed.): *The Cambridge Guide to Second Language Assessment*, New York: Cambridge University Press, 30-36.
- Alderson, J. Charles (1996): "Do corpora have a role in language assessment?", in: Thomas, Jenny / Leech, Geoffrey (Ed.): *Using corpora for language research: studies in the honour of Geoffrey Leech*, London: Longman, 248-259.
- Barker, Fiona (2006): "Corpora and language assessment: trends and prospects," in *Research Notes* 26, 2-4.
- Barker, Fiona (2010): "How can corpora be used in language testing?", in: O'Keeffe, Anne / McCarthy, Michael (Ed.): *The Routledge Handbook of Corpus Linguistics*, London: Routledge, 633-645.
- He, Lianzhen / Dai, Ying (2006): "A corpus-based investigation into the validity of the CET-SET group discussion", in *Language Testing* 23, 3, 370-401.
- Karimi, Neda (2011): "C-test and vocabulary knowledge", in: *Language Testing in Asia* 1, 4, 46-58.
- Lienert, Gustav A. / Raatz, Ulrich (1998): *Testaufbau und Testanalyse*, Weinheim: Psychologie Verlags Union.
- Messick, Samuel (1989): "Validity", in: Linn, Robert L. (Ed.): *Educational measurement*, Washington, DC: American Council on Education and National Council on Measurement in Education, 13-103.
- O'Sullivan, Barry (2012): "Assessment issues in languages for specific purposes," in: *The Modern Language Journal* 96, Focus Issue, 71-88.

How to find Writers of Vegan Restaurant Descriptions

- Sinclair, John (1991): *Corpus, concordance, and collocation*, Oxford: Oxford University Press.
- Taylor, Linda B. et al. (2003): "Exploring the relationship between language corpora and language testing," in: *Symposium at 25th Language Testing Research Colloquium*, 22-25.