

THE EFFECT OF NATIVE ENGLISH ACCENTS ON COMPREHENSIBILITY AMONG DUTCH
STUDENTS

by

Annabel Julia Schmitz

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE
OF BACHELOR OF ARTS

English Language and Culture

Utrecht University, 2015

Supervisors: Marco van de Ven (Utrecht University) & Elena Tribushinna (Utrecht University)

June 2015

Abstract

This thesis aims to find whether there are significant differences between the comprehension of native English accents among young Dutch students of English and if familiarity with these accents might influence comprehensibility. Many researchers found that a speaker's accent has effect on listening assessment scores (Eisentein and Berkowitz, 1981; Anderson-Hsieh and Koehler, 1988; Bilbow, 1989; Derwing and Munro, 1997) and that there might be a connection between a degree of familiarity and comprehensibility of an accent (Gass and Varonis, 1984; Bradlow and Brent, 2008; Adank, Evans, Stuart-Smith and Scott, 2009; Okay and French 2014). Three native English accents were included in the test, namely British, American and Australian. By means of a listening test, corresponding to the students' listening level, and a survey, students' listening comprehension and familiarity with the accents were assessed. The listening test included 15 multiple choice questions and the survey consisted of four Likert scale questions. Findings include: (1) No significant differences between the assessment scores were found, each accent was equally comprehensible; (2) neither male nor female students performed significantly different; (3) there was no correlation found between the familiarity and comprehensibility of an accent.

Contents

1. INTRODUCTION	1
2. THEORETICAL FRAMEWORK	1
2.1 FAMILIARITY OF ACCENT	2
2.1 NON-NATIVE ENGLISH LISTENING COMPREHENSIBILITY	2
2.2 NATIVE ENGLISH ACCENT COMPREHENSIBILITY.....	3
3. METHOD	6
3.1 DETERMINING COMPREHENSION LEVEL.....	6
3.2 SELECTING SPEAKERS.....	7
3.4 TEST DESIGN.....	8
3.4.1 <i>Test and survey choices</i>	9
3.4.2 <i>Scoring methodology</i>	10
3.5 PARTICIPANTS.....	11
4. RESULTS	11
4.1 MULTIPLE CHOICE TEST RESULTS.....	11
4.2 SURVEY FINDINGS.....	14
4.3 COMPARATIVE FINDINGS	15
5. DISCUSSION	16
5.1 LIMITATIONS.....	18
6. CONCLUSION	20
REFERENCES	20
APPENDICES	23

1. Introduction

Within the vast and ever expanding English speaking world a tremendous amount of English accents are spoken and even more accents emerge and are officially recognized by scholars and linguists. The immense number of accents has implications for English language teaching (ELT) and how language testing might have to deal with such an extensive amount of accents. The expansion means the demand for multidialectal listening skills for ELT is becoming increasingly prevalent. To obtain such multidialectal listening skills, ELT learners will have to be confronted by diverse accents for a comprehension of a variety of accents might be necessary to understand and communicate in different English speaking contexts. Goh (1999) points out that a majority of listeners find the accent of a speaker the predominant factor influencing listening comprehension. For this reason it is important to look at how various accents influence comprehensibility among second language learners and how this knowledge should be incorporated in listening tests.

2. Theoretical Framework

Many scholars have expressed concern about, and questioned, the use of only a single English accent for assessing second language listening comprehension (L2 LC) (Harding 2011; Taylor and Gernanpayeh 2011; Abeywickrama 2013; Ockey and French 2014). The predominant argument is the changing demographics within the English speaking world causing a need for multiple varieties to be included in listening comprehension tests. These tests would then reflect more accurately how well the participants might be able to function in diverse English contexts. However, an abundant body of research supports the claim that a speaker's native as well as non-native accent can affect listening comprehension significantly enough to affect assessment scores (e.g. Eisentein and Berkowitz, 1981; Anderson-Hsieh and Koehler, 1988; Bilbow, 1989; Derwing and Munro, 1997). Therefore, if multiple accents are included in

listening comprehension tests it can affect fairness of such tests because certain accents might be naturally more difficult to comprehend than others. Under this hypothesis, choice of accent might not only influence test results but also differ based on the test taker, see Taylor (2006).

2.1 Familiarity of accent

A major influence on the comprehensibility of diverse accents is found to be the participant's familiarity with a particular accent (Gass and Varonis, 1984; Bradlow and Brent, 2008; Adank, et al., 2009; Okay and French 2014), although different claims as to how the familiarity effect interacts with comprehensibility have been made. Familiarity is defined as the amount of experience a listener has with a particular variety, and naturally listeners have much experience with the non-native variety based on their first language. Consequently, the same L1 criterion is a subset of the familiarity criterion. Tauroza and Luk (1997) found that whether a particular accent causes listening comprehension difficulty for L2 learners of English is predominantly based on the degree of familiarity of the accent. Their study amongst Hong Kong high school students even suggested that whether the speaker's accent is similar to the listener's accent is a secondary issue.

2.1 Non-native English listening comprehensibility

There is a considerable amount of discussion about the issue of including non-native accents in listening comprehension tests because the population of non-native speakers of English exceeds that of native speakers (Yano, 2001). Researchers such as Major et al. (2002) suggested that a shared L1 could influence comprehensibility. A lot of research has shown that the performance of students is significantly better when listening to a non-native English accent similar to their L1 (Ekong, 1982; Bent and Bradlow, 2003; Moinzadeh, Rezaei, and Dezhara, 2012). Recent studies by Sadeghi and Zeinali (2014) and Abeywickrama (2013) not only found

that test takers who had listened to a non-native accent performed better but moreover preferred the non-native English accent over a native English accent. Although there are many studies that support including non-native accents without creating significant unfairness in tests, the contradictory results should also be highlighted. The studies conducted by Smith and Bisazza (1982) and Eisentein and Berkowitz (1981) found that standard British and American speakers were more comprehensible for L2 listeners than foreign accented speakers. In addition, studies supporting the inclusion of non-native accents solely looked at non-native varieties corresponding to or similar to the listener's L1.

2.2 Native English accent comprehensibility

Another interesting discussion exists within the spectrum of native varieties. Between these native varieties great differences exist in terms of accent, e.g. British, American, South African. To what extent these differences affect the listening comprehension has been researched by various scholars (e.g. Matsuura, Chiba and Fujieda, 1999; Major et al., 2005; Adank, et al., 2009).

Matsuura, et al., (1999) studied intelligibility and comprehensibility among Japanese university L2 students for both familiar and unfamiliar English accents, using native American and Irish English speakers. The participants heard a familiar accent, the dominant accent encountered in school setting namely American, and an unfamiliar accent, an uncommon variety namely Irish. The students perceived the speaker with a familiar accent easier to comprehend; however, the superior comprehension did not lead to improved test results.

A study conducted by Major et al. (2005) compared regional, ethnic and international dialects of English with Standard American and found that for ESL listeners there was a significant effect on comprehensibility when hearing ethnic and international dialects. ESL

listeners found these dialects more difficult to perceive. However, no effect was observed for the regional dialects and Major et al. therefore suggest these should be used in listening comprehension tests.

Important research was conducted by Ockey and French (2014). They compared the comprehension of one US, four Australian, and four British English speakers among TOEFL iBT test takers. Most importantly, they also measured and considered the strength as well as the familiarity with an accent in their assessment. They note that some of the previous studies into the effect of accents on listening comprehensibility had partially failed to do so (p.3). They address the issue that the mixed results within this field of study are probably because there is no clearly defined familiarity threshold beyond which familiarity has effect on L2 listening comprehension. Ockey and French tried to investigate such a threshold and conclude that the strength of an accent and familiarity both affect listening comprehension, even when quite light accents are used (p.20).

Interestingly most of the studies into the effect of accent on listening comprehension have been conducted among academic demographic, e.g. enrolled or future university students. Ockey and French's research suggests that accent does matter for older and more advanced students, for the TOEFL iBT test is used to measure an academic level of English proficiency. Their findings incite an interesting question, what would the impact of accent be on the listening comprehension of much younger and beginner L2 learners of English?

In the Netherlands, English is the predominant second language taught in every secondary school and is compulsory for all students regardless of their educational level. Nowadays, English is a subject that requires a pass for students to even advance to the next year. In addition to their English education at secondary school, Dutch children are also taught the basics of English in primary school, although not as frequent as in secondary school. This was implemented by the government in 1986. English has therefore a prominent

role within the Dutch educational system and all students are required to master the accompanying skills to a certain degree, including listening comprehension.

A Dutch student's listening comprehension naturally develops and improves itself throughout secondary school; however, the assessment of this skill starts very early on in the Netherlands namely at their first year of secondary school. These listening tests usually involve a small variety of native English accents, whilst the amount of accents included gradually expands towards the latter years. These young learners are especially interesting for they have only been exposed to a select variety of mostly native English accents but might or might not already show differences in the comprehension of such accents. Whether they are familiar with some or any of these accents is an interesting relatable question which could possibly affect their listening test scores and might have implications for future multidialectal test design. To shed light on the situation in the Netherlands, this study aims to answer the following research questions in regard to assessing accents in L2 listening comprehension:

RQ: Is there a difference between the listening comprehension of the British, American and Australian accent among Dutch adolescent secondary school L2 students of English?

RQ_{sub}: If so, is part of the difference relatable to the familiarity with these accents?

3. Method

3.1 Determining comprehension level

To establish what the listening proficiency level of second year students of University preparatory education (VWO) is, currently applied listening comprehension tests are used as a starting point. One of these listening comprehension tests is the central final examination listening test from 2012 which was designed by the Netherlands Institute for Educational Measurement, from here on referred to as Cito. The listening proficiency of all the participants in this study is tested with the central final examination listening test of the pre-vocational secondary education, Dutch VMBO, which is normally taken after four years of secondary school.

The relevance of the Cito listening test lies in its correlation to the CEFR levels. The Common European Framework of Reference: Learning, Teaching, Assessment, abbreviated as CEFR, is an important framework for modern language education in Europe. Created by the Council of Europe as a result of over twenty years of research into language proficiency and as of 2001 recommended by the European Council Resolution to validate language ability, it functions as a method of assessing, teaching and learning L2 languages (Council of Europe, 1-8). The CEFR's framework exists of common reference levels and distinguishes three levels of foreign language proficiency: A, basic language user; B, independent language user and C; proficient language user. In practice each level is divided into a lower and higher degree of proficiency thus creating six levels in total, ranging from *Breakthrough A1* to *Mastery C2*. Each level is defined according to a scale, provided with definitions in the form of *can do statements*. For example a student with A2 listening proficiency “[c]an understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment) (Council of Europe 24).”

Most Dutch schools make use of the CEFR proficiency standards, although this is not mandatory, but most importantly these proficiency standards can be applied to Cito's designed listening comprehension tests. The final examination listening test of VMBO can indicate whether the listening comprehension of students is located at an average of CEFR level B1: Threshold.

3.2 Selecting speakers

Three native speakers of English voiced the sound fragments used in the study. All speakers are currently bilingual teachers on a Dutch secondary school. The speakers were given time to familiarise themselves with the text before recording and guided by the researcher during recording. The speakers' nationality, gender, age, area/city of origin and their length of stay in the Netherlands is summarised in Table 1. To make the strength of accent variable as consistent as possible this study included light accents which are close to the standard dialect of their native country, namely Britain, America and Australia, but still significantly distinctive from each other.

Table 1: Details of speakers

#	Birth Nationality	Gender	Age	Accent	City/Area of Origin	Years of Stay in Netherlands
Sp. 1	American	Female	34	Inland Northern American	Holland, Michigan	9
Sp. 2	British	Female	54	West Midlands	Burntwood, Staffordshire	23
Sp. 3	Australian	Male	36	South Australian	Adelaide, South Australia	15

3.4 Test Design

The test consists of two parts: a multiple choice test to measure the comprehension of each accent and a survey to give insight into the familiarity of the accents, see Appendix A.

The multiple choice test consists of 15 questions about a total of 15 sound fragments, of which five are voiced in British, five in American and five in Australian, in random order. These sound fragments are based on excerpts from the transcripts from the B2 listening proficiency section from the British Council website, see Appendix B (Listening Skills Practice Level). Three native speakers of English each voiced 15 clips in total. To remove any possible correlation between the contents of a sound fragment and its comprehensibility all fifteen sound fragments are recorded in every accent. These 45 fragments are then used to create three multiple choice tests so every fragment is voiced in each accent exactly once. The fragments are on average between 30 and 40 seconds long. Students have 10 seconds after each fragment to answer a multiple choice question which requires them to interpret the information given by the speaker. Each multiple choice question contains four answers.

Insight into the degree of familiarity of the British, American and Australian accent among the students is obtained through an accent familiarity survey. Again students hear three short sound fragments, one for every accent, followed by four questions each. These questions are not related to the contents of the text but purely to the speaker's accent.

First off, the students have to answer a question regarding the clarity of the speaker. The second and third questions are related to the student's experience with each accent, the indirect familiarity rates how often a participant heard any of the accents on TV, radio or internet and the direct familiarity rates the experience with each accent in face to face communication. A 5 point Likert scale is used to answer the questions (see Appendix A). Lastly, students have to determine which accent they think the speaker has by writing down the speaker's country.

3.4.1 Test and survey choices

As mentioned above the listening proficiency level of adolescent Dutch students correlates to the average CEFR level B1: Threshold. However, students' Cito listening tests results are relatively high, an average of 69% amongst all participants. These high grades are undesirable for this study since it is essential to observe significant differences if there are any. The test should be hard enough so that a small difference in comprehensibility can be easily measured in percentages. Ideally the fragments should be challengingly comprehensible. The aim is to create a test where the average student has around 50% correct. To optimise the statistical power of the test and to create distinctiveness amongst the participants a higher CEFR level will therefore be used, namely B2: Vantage. The listening test designed for this study conforms to a fixed set of criteria based on the CEFR B2 level descriptors and domains for listening, thus creating a clear and justified framework eliminating as many possible variables. The overall description of B2 reads:

- [A student c]an understand standard spoken language, live or broadcast, on both familiar and unfamiliar topics normally encountered in personal, social, academic or vocational life. Only extreme background noise, inadequate discourse structure and/or idiomatic usage influence the ability to understand.
- [A student c]an understand the main ideas of propositionally and linguistically complex speech on both concrete and abstract topics delivered in a standard dialect, including technical discussions in his/her field of specialisation.
- [A student c]an follow extended speech and complex lines of argument provided the topic is reasonably familiar, and the direction of the talk is sign-posted by explicit markers (Overview CEFR Scales 8).

Within the B2 listening proficiency level the following subskills, i.e. global descriptors, are also defined:

Understanding interaction between native speakers:

- *Can keep up with an animated conversation between native speakers.*
- *Can with some effort catch much of what is said around him/her, but may find it difficult to participate effectively in discussion with several native speakers who do not modify their language in any way.*

Listening as a member of a live audience

- *Can follow the essentials of lectures, talks and reports and other forms of academic/professional presentation which are propositionally and linguistically complex.*

Listening to announcements & instructions

- *Can understand announcements and messages on concrete and abstract topics spoken in standard dialect at normal speed.*

Listening to media audio & recordings

- *Can understand recordings in standard dialect likely to be encountered in social, professional or academic life and identify speaker viewpoints and attitudes as well as the information content.*
- *Can understand most radio documentaries and most other recorded or broadcast audio material delivered in standard dialect and can identify the speaker's mood, tone etc. (Overview CEFR Scales 8-9)*

3.4.2 Scoring methodology

Guessing strategies are taken into account in the multiple choice test. Holt's formula for calibrating the marking scheme and establishing a fair penalty to discourage guessing has been used (2006). a is defined to be the point value of a correct answer and θ the number of answers in a multiple choice question. Then the 'fair' value for a wrong answer b_{FAIR} in points to be deducted is given by:

$$(1) b_{FAIR} = \frac{a}{\theta - 1}$$

For this study a b_{FAIR} of 1 is desirable, with $\theta=4$ for there are four answers per question, which gives a equals 3. Thus three points are rewarded for a correct answer, no points for not filling in the question and a one point deduction for a wrong answer. In this way it is not valuable for the participants to guess an answer, a participant who answers everything randomly will end up with an expected score of 0. By using this formula the minimum amount of points is a total of -15 and the maximum amount is 45. The result is scaled by dividing this value by the maximum amount obtainable to get a value ranging from -0.33 till 1. This will be used as the grading scale throughout the rest of this study.

3.5 Participants

This study includes 78 Dutch secondary school students that are in their second year of university preparatory education, Dutch VWO. The students are normally divided into three classes with an average of about 25-30 students per class. All students have Dutch as a L1 and English as L2. The students' ages range from 12 to 14 years old. Additional demographic information is shown in Table 2. Their current English teachers base their accents on a British model of pronunciation. Young Dutch children have been frequently exposed to British and American varieties, in school settings and via (social)media, but often have not had similar contact with other native varieties such as Australian.

Table 2: Demographic details

Group	N	Sex		Age	
		M	F	Mean	St.Dev
I	28	9	19	13.64	0.49
II	24	12	12	13.71	0.55
III	26	11	15	13.46	1.98
Total	78	32	46	13.60	1.21

4. Results

4.1 Multiple choice test results

The general test results from the three classes combined are shown in Table 3 and Figures 1 to 4 below. These general results show that the subject scored best on those questions that were voiced by an American accent, although it is apparent that the results are also more varied for this particular subset of data. On average the subjects scored 0.37 points which is significantly higher than a random answer sample, this concludes that the students could at least answer a part of the questions but could not answer every question correctly. The study did overshoot its target difficulty, 50%.

Table 3: Average score and (standard deviations) per accent

Accent	Average Score	St. dev.
American	0.43	0.36
British	0.35	0.32
Australian	0.35	0.29
Total	0.38	0.20

Figure 1

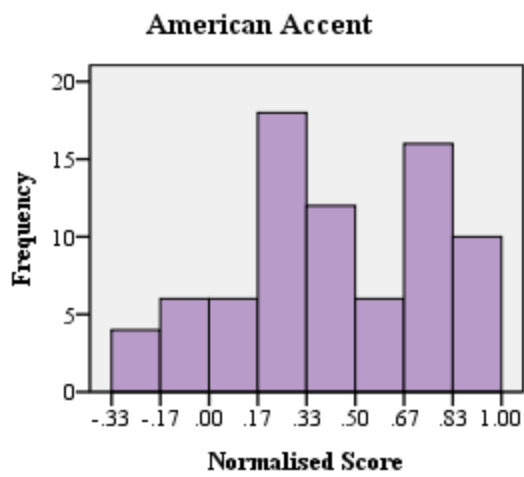


Figure 2

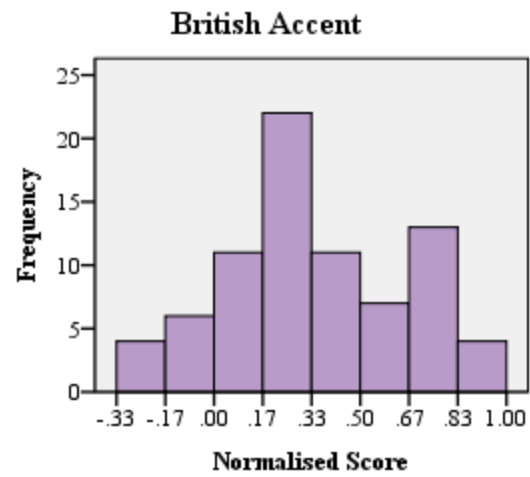


Figure 3

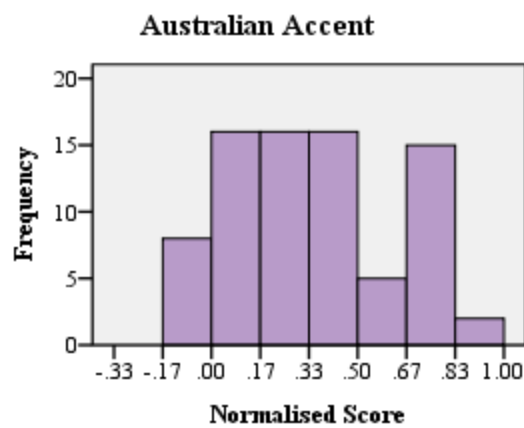
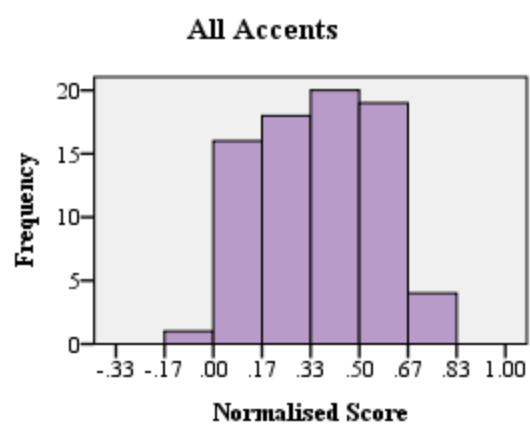


Figure 4



In Table 4 the differences between the male and female scores are shown per accent as well as for the total score. As can be seen from the p-values no significant differences were found for all p-values > 0.05 . Interestingly, male and female participants performed equally.

Table 4: Means, St.Dev, t-values and p-values per accent and total for male and female

	American	British	Australian	Total
Male	0.43	0.37	0.41	0.40
St. dev.	0.36	0.35	0.29	0.20
Female	0.43	0.34	0.31	0.36
St. dev.	0.37	0.30	0.29	0.20
t-values	0	0.38	1.48	0.91
p-values	$> .999$.71	.14	.37

To see whether the observed differences between the accents are significant a one way ANOVA test was conducted, using a Tukey analyses, see Table 5. No significant differences were observed between the three accents within the chosen confidence intervals which were selected for $\alpha=0.05$.

Table 5: Overview One way ANOVA all accents

ANOVA Results					Tukey HSD	
Type I	Type II	Mean Difference (I- II)	Std. Error	p-value	95% Confidence Interval	
					Lower Bound	Upper Bound
American	British	0.081	0.052	0.268	-0.042	0.204
	Australian	0.081	0.052	0.268	-0.042	0.204
British	Australian	0.000	0.052	1.000	-0.123	0.123

As an example, suppose a listening test was conducted using this study's scoring method a student that would be awarded a score of 7 out of 10 could have anywhere between a 5.8 and a 8.2 depending on the accent.

4.2 Survey findings

Table 6 shows the results of the survey conducted to measure clarity, direct familiarity (real) and indirect familiarity (media) for each accent. Values that significantly differ from a neutral result of 3 on the Likert scale are shown in **bold**, for positive, and *italics*, for negative. The student t-test with $k=78$ gives $z=1.990$, this z value was in turn used for a categorisation into positive, negative and indeterminate results. As can be seen in Table 6, survey results were overwhelmingly positive for clarity of accent. Participants showed to have plenty of experience with the British accent in media, while real experience with the American and Australian accent is rare. Almost no participants filled in any questions with a rating of 1 on the Likert scale for the clarity and media questions. The last question, where students were required to identify the speaker's country of origin, posed several difficulties. Retrospectively the question turned out to be awkwardly phrased and therefore students misinterpreted the question. The question has been discarded from the results.

Table 6: Survey results binned by Likert scale rating.

		Rating Frequency					Descriptors	
		1	2	3	4	5	Average	St.Dev
American	Clarity	1	13	26	28	10	3.42	0.96
	Media	4	18	29	17	10	3.14	1.08
	Real	20	28	17	9	4	2.35	1.14
British	Clarity	0	7	25	30	16	3.71	0.90
	Media	0	10	27	29	12	3.55	0.91
	Real	12	17	28	17	4	2.79	1.11
Australian	Clarity	1	14	21	30	12	3.49	1.00
	Media	3	16	34	20	5	3.10	0.93
	Real	9	26	25	14	4	2.72	1.06

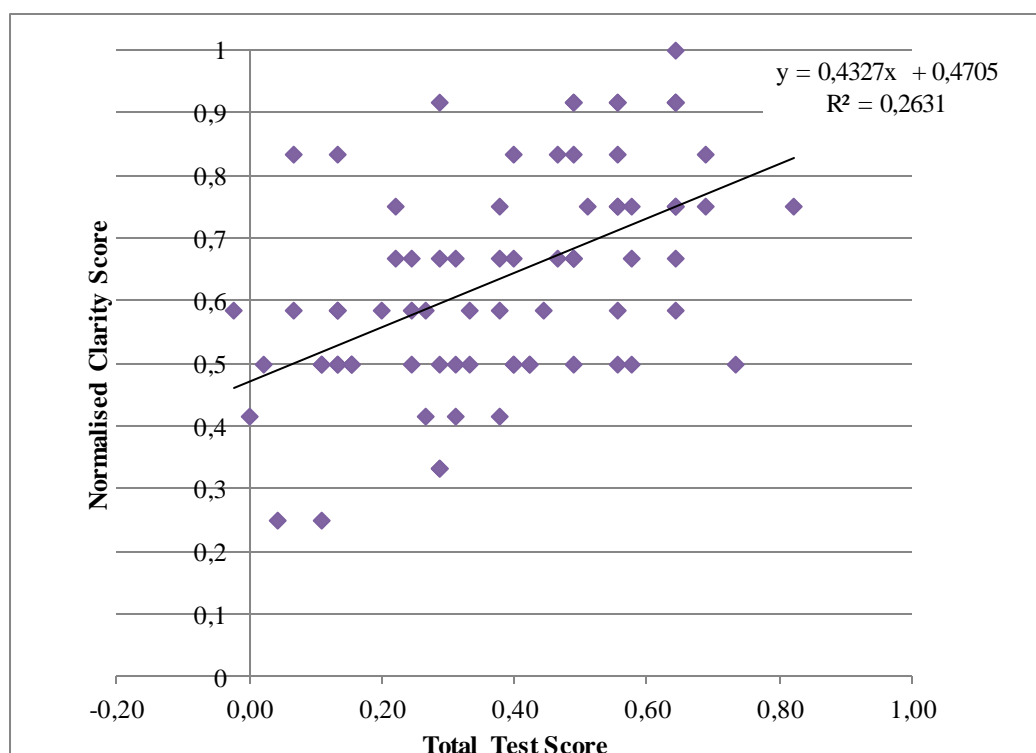
4.3 Comparative findings

As a control experiment this study compared the perceived clarity of the speaker by the participants with the test results. In order to calculate a normalised value ranging from 0 to 1 for the clarity evaluation questions, the following formula was used:

$$S = \frac{C_1 + C_2 + C_3}{12}$$

Where S is the normalised score and the C variables are the Likert scale results for the three accents, in no particular order. Compare the value S for each of the total score per participant yields the following scatterplot, see Figure 5. This scatterplot can infer the usefulness of the survey. In addition the high average on the y-axis compared to the x-axis indicates that the clarity questions were answered with high marks overall, see also Table 6, resulting in a positive correlation. Assuming a strong negative correlation between strength of accent and clarity shows that the strength variable in this study is quite low.

Figure 5: Scatterplot showing clarity versus total test score



To see whether there was any correlation between the test scores of the particular accents and the familiarity, a paired t-test was used. For the familiarity measure the maximum of the two familiarity questions, direct familiarity and indirect familiarity, was used. In all cases there was no significant difference between these variables for all H_0 hypotheses could not be rejected based on their corresponding p values, respectively 0.741, 0.299 and 0.856 for the American, British and Australian accents.

Table 7: Correlation between test result and familiarity measure

Pair	N	Correlation	p-value
American	78	0.038	0.741
British	78	0.119	0.299
Australian	78	0.021	0.856
Total	78	0.190	0.095

5. Discussion

In this study no significant difference in test scores between any combination of two accents from those analysed was found, see Table 3, nor were there any significant differences between male and female test scores, see Table 4. A possible reason why no significant differences could be observed might be due to the strength of accent variable, which was quite low by design in study. This variable was not officially measured; however, the method of speaker selection used implies the former. As Ockey and French (2014) have suggested accents with a very low strength variable, those that would scale lower than a two on a Likert scale by a panel of judges, would be unlikely to show any significant impact on listening comprehension scores (19). Note that this result is independent of accent familiarity.

The survey results show clearly that participants were most familiar with the British speaker. The British speaker consistently has the highest scores for all three variables. The United Kingdom is geographically close and British English is the preferred reference accent in Dutch classrooms (Van der Haagen, 1993, As cited in Koet, 2007). The difference in scores between the American and Australian direct familiarity are contrary to expectations. The relatively higher direct familiarity score for Australian seems strange at first because Australia is a less populated country and students are less likely to be exposed to the Australian accent than the British or American one. The American accent has a prominent role in popular culture and there are simply more Americans than Australians. However, a small subset of the participants actually encountered the very speaker as a teacher in their first half year of secondary school. These participants may (un)consciously have given a higher direct familiarity score because they might have recognised the speaker as a person.

In addition, this study found no correlation between perceived familiarity and comprehension which should exist according to Gass and Varonis (1984), Bradlow and Brent (2008), Adank, et al. (2009). Little evidence of a strong correlation was found. To refer back to Ockey and French, their study implied that a correlation between familiarity and comprehension does exist but depends on strength of accent being sufficiently high. There is reason to believe that this study does not meet the strength requirements to be able to measure this correlation. For future research it might be interesting to compare stronger and more rural accents to see more pronounced results.

Seeing that there are no significant differences between the test scores, all of the accents used in this study could justifiably be used on listening tests without creating any unfairness. The accents used in this study are all categorised as Inner Circle varieties of English in Braj Kachru's model of World Englishes (1982). A possibility exists that different results could be found if more Inner Circle varieties are included, e.g. Irish, South African,

Canadian, Caribbean and New Zealand English, since young Dutch students are probably less familiar with these accents. It would also be interesting to look at non-native varieties since much research has been done into the effects of including non-native varieties and the attitudes towards these varieties. In the Netherlands it is very rare to encounter a non-native variety in official listening tests in the first two years of secondary school. However, the inclusion of such varieties might show different test results, especially if the accent is based on their L1, in other words Dutch-English (i.e. Denglish), see Ekong (1982), Bent and Bradlow (2003), Moïnzadeh, Rezaei, and Dezhara (2012).

5.1 Limitations

The first and foremost limitation in practical consideration is the size of the test group. With $N=78$ it becomes rather difficult to find the type of small difference usually observed in this kind of study, see also Ockey and French. Participants, likely due to being young thus unreliable and of greatly varying English L2 level, cause results to have a large statistical variation. To get good estimators for quantified measurements of variables such as comprehension and familiarity a larger set of results is necessary for the same amount of precision.

This study used a limited set of speakers who only have a very select accent. No formal study was conducted on the validity of the accent of the speakers. The speakers were chosen based on a close resemblance to the standard dialect spoken in the related country but based on a subjective opinion of the researcher. If it is truly desirable to measure listening proficiency in hearing a typical native speaker it becomes necessary to include a range of accents in the particular country, ideally weighed with the proper distribution of the dialects in the country.

The ability to comprehend accents and dialects can vary based on the gender of the speaker. This study included two female speakers and one male speaker while preferably speaker would be one gender only to homogenise the result more optimally. There could be a hidden influence due to one of the speaker being male. In the ideal situation the sound fragments used feature the proper gender distribution; however, this would require a large amount of speakers just like the conclusion of the previous paragraph.

The final limitation lies in test and survey questions. Each student can obviously only take one of the three test variations. These test variations were designed to eliminate the effects of question difficulty aside from comprehension based on accent on our results. Particular traits of a question can be eliminated by asking each question in each accent. There is a silent assumption being made here: a student taking a particular test is not statistically expected to have a disparate skill level, e.g. just like a random sample. However, groups of students were handed out each of the test variations by class. This might invalidate the assumption because classes are predetermined and may not contain a homogenized random selection. Particular classes may score higher on tests in general, based on for example the English teacher, or student natural intelligence. For example, five students of one of the classes followed a gymnasium programme but resided within the VWO class. Suppose one of the questions was generally more difficult and therefore required a higher cognitive ability to answer correctly, then the average score for the accent in which this question was asked to the more proficient group would be higher. This limitation could be prevented by randomly assigning tests to a large group of participants. The survey questions measure the subjective judgement of the participants. There might be a large amount of variation between participants perception, furthermore it is not possible to quantify the results absolutely so they are only useful relatively and within the same study.

6. Conclusion

For the American, British and Australian accent no effect on comprehensibility among Dutch adolescent secondary school L2 students of English has been found. For neither female nor male participants any significant effect was found. This means all these accents are suitable to include in English listening tests without creating any unfairness. In addition, there was no connection between test scores and familiarity with any of the accents, which indicates that either familiarity has no effect on test results or that it is necessary to include stronger accents to observe any effect. More research is necessary to provide a clearer image on this topic.

References

- Abeywickrama, P. (2013). Why not Non-Native Varieties of English as Listening Comprehension Test Input? *RELC Journal* 44(1), 59–74.
- Anderson-Hsieh, J. and K. Kohler. (1988) ‘The Effect of Foreign Accent and Speaking Rate on Native Speaker Comprehension,’ *Language Learning*, 38(4), 561–613.
- Bent, T., & Bradlow, A. R. (2003). The Interlanguage Speech Intelligibility Benefit. *Journal of the Acoustical Society of America*, 114(3), 1600–1610.
- Cito (2007). *De koppeling van de Cito kijk- en luistertoetsen moderne vreemde talen aan het Europees Referentiekader*. Arnhem: Cito.
- Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press. Web. 19 May, 2015.
- Council of Europe (2001). *Overview of all CEFR scales*.
http://www.coe.int/t/dg4/education/elp/elp-reg/Source/Key_reference/Overview_CEFRscales_EN.pdf. Web. 19 May, 2015.

- Derwing, T. M. and M. J. Munro. (1997). 'Accent, Intelligibility, and Comprehensibility: Evidence from Four L1s,' *Studies in Second Language Acquisition* 20, 1–16.
- Ekong, Pamela. (1982). On the Use of an Indigenous Model for Teaching English in Nigeria. *World Language English*, 1(2), 87-92.
- Gass, S. and E. M. Varonis. (1984) 'The Effect of Familiarity on the Comprehensibility of Non-Native speech,' *Language Learning* 34(1), 65–89.
- Harding, L. (2011). Accent and Listening Assessment: A Validation of the Use of Speakers with L2 Accents on an Academic English Listening Test. Peter Lang.
- Holt, A (2006): An Analysis of Negative Marking in Multiple-Choice Assessment. *19th Annual Conference of the National Advisory Committee on Computing Qualifications* (NACCQ 2006), Wellington, New Zealand, 115-118.
- Kachru, B. (1982). *The Other tongue: English across cultures*. Urbana: University of Illinois Press.
- Koet, A. (2007). *Polder English in Dutch ears: Empirical studies on the evaluation of the pronunciation of English as a foreign language*. PhD dissertation, University of Amsterdam.
- Listening Skills Practice Level B2. (n.d.). British Council. Retrieved May 22, 2015, from <http://learnenglishteens.britishcouncil.org/skills/listening-skills-practice>
- Major, R., S. Fitzmaurice, F. Bunta, and C. Balasubramanian. (2002). 'The Effects of Non-Native Accents on Listening Comprehension: Implications for ESL Assessment,' *TESOL Quarterly* 36(2), 173–90.
- Major, R., Fitzmaurice, S., Bunta, F., & Balasubramanian, C. (2005). Testing the Effects of Regional, Ethnic, and International Dialects of English on Listening Comprehension. *Language Learning*, 55(1), 37-69.

- Matsuura, H., Chiba, R., & Fujieda, M. (n.d.). Intelligibility and Comprehensibility of American and Irish Englishes in Japan. *World Englishes*, 18(1), 49-62.
- Moinzadeh, A., Rezaei, O., & Dezhara, S. (2012). The Effect of Non-Native Accent on Iranian EFL Learners' Listening Comprehension, Focusing on Persian Accent of English. *Journal of Language Teaching and Research*, 3(5), 967-972.
- Ockey, J. G. and French R. (2014). From One to Multiple Accents on a Test of L2 Listening Comprehension." *Applied Linguistics Advance Access*, 1-24 . Web.
doi:10.1093/applin/amu060
- Sadeghi, K., & Zeinali, M. (2014). The Effect of Input Source (Native vs. Non- native) and EFL Learners' Perceptions Towards It, on Their Listening Performances, across Gender. *IJRSLL International Journal of Research Studies in Language Learning*, 3(3), 45-54.
- Smith, L. and J. Bisazza. (1982) 'The Comprehensibility of Three Varieties of English for College Students in Seven Countries,' *Language Learning* 32 (2), 259–69.
- Tauroza, S. and J. Luk. (1997). Accent and Second Language Listening Comprehension. *RELC Journal* 28, 54–71.
- Taylor, L. 2006. 'The Changing Landscape of English: Implications for English Language Assessment,' *ELT Journal* 60(1), 51–60.
- Taylor, L. and A. Garenpayeh. (2011). Assessing English for academic purposes: Defining and operationalizing the test construct, *Journal of English for Academic Purposes* 10, 89–101.
- Yano, Y. (2001). World Englishes in 2000 and beyond. *World English*, 20 (2), 119–31.

Appendices

Appendix A

Let's Listen

Bij deze luisteropdracht ga je luisteren naar 15 korte fragmentjes en beantwoord je steeds een multiple choice vraag. Er zijn pauzes.

LET OP! Het is heel belangrijk dat je niet gaat gokken! Als je het niet weet vul de vraag dan niet in (anders krijg je minpunten).

The following speakers talk about their heroes.

answer

1. *What was so special about Mary Anning?*

- a. She discovered the largest dinosaur ever found in England.
- b. She had no formal education; she was a self-taught archeologist (archeoloog).
- c. She became the first female scientist to working alongside male scientists.
- d. She discovered all dinosaurs had died before 1811

2. *Kalash Satyarthi.....*

- a. has a factory with children that make rugs (tapijten).
- b. is born in South Asia, Nepal, and has no education.
- c. sometimes risks his life saving enslaved children.
- d. saved over eight million children from abuse.

3. *Why did Rachel Carson write Silent Spring in 1962?*

- a. To show how humans endanger the ecosystem by using chemicals
- b. To promote DDT chemicals.
- c. To point out that nature is in danger, especially the sea.
- d. To explain that nature would be better off without humans.

4. *This speaker's hero is John Lennon because:*

- a. His message about peace for everyone was very important
- b. He wrote a beautiful song called *Immersion* in 1971.
- c. He was very popular at the time.
- d. He was an amazingly easy, cheesy and serious person.

The following speakers give tips on how to study.

5. *What should you do when studying according to this tip?*

- a. Make a strict timetable until the exam and do exactly what it says.
- b. Make a study plan for the time you intend to study, but adjust it.
- c. You need to plan at least six days ahead.
- d. You need a timetable, a schedule as well as a plan.

6. *What does this tip say about your studying environment (omgeving)?*

- a. It's important to have absolutely no noise in the background.
- b. The TV should be on the whole time when studying.
- c. It's your own choice to have background noise.
- d. If you want to concentrate well, you should listen to music.

7. *According to this study tip, what **shouldn't** you do.*

- a. Treat yourself in study breaks with Coke and M&Ms.
- b. Divide your studying periods by using breaks.
- c. Study in shorter periods so your brain works more efficient.
- d. Adjust your breaks according to your preference.

8. *What does this tip say about how you should study?*

- a. You should only take notes if you want to.
- b. People all study differently, but most people just read textbooks.
- c. You must use different methods of studying.
- d. It is very important to take notes when studying.

9. *What is said about mind maps?*

- a. Writing down random sentences is better than making a mind map.
- b. Mind maps are better if details are unimportant
- c. A combination of summaries, notes and mind maps is best.
- d. Mind maps are similar to our brains' thinking process.

The following speakers talk about dream jobs in Australia

10. *What is **NOT** a thing an "outback adventurer" does?*

- a. A range of different activities, such as sleeping outside.
- b. Eating foreign food, and maybe strange insects
- c. Getting to know about aboriginal culture and their traditions
- d. Discovering several things elderly people could do on their (work) holidays

11. *Where do Queensland park rangers patrol?*

- a. Mainly along the coast and underwater.
- b. Around Lizard Island and the Great Barrier reefs.
- c. In a very wide park area, on land but not near the sea.
- d. In a dream world.

12. *If you want to be a wildlife caretaker, which quality is **most** important*

- a. Showing interest in wildlife and being able to handle animals
- b. Having the ability to fight with sharks
- c. Being able to cycle, kayak and use a jeep for transport
- d. Making sure people don't pollute (vervuilen) beaches.

13. *What will you **NOT** do as a lifestyle photojournalist?*

- a. Take pictures of the newest, most interesting cafés.
- b. Visit and write about cool musical festivals.
- c. Keep a day to day journal about wildlife.
- d. Write reports on tourist activities in Victoria.

14. *The job of a taste master is to*

- a. Taste all kinds of food in and around Western Austria.
- b. Have experience with different types of cooking.
- c. Promotes all different kinds of establishments in the catering industry (horeca).
- d. Has to like every type of drink and food, especially seafood.

15. *As a "chief funster", what is **NOT** part of your job?*

- a. Using social media to inform the public about popular events.
- b. Being very interested in Australian sports.
- c. Being involved in promoting the firework spectacle in Sydney Harbour.
- d. Working on different entertaining festivals and events.

Survey

Je krijgt weer drie korte fragmentjes te horen. Het gaat hier alleen om zijn/haar manier van spreken. Beantwoord de vier vragen hieronder. - - betekend helemaal niet/nooit en + + betekend heel erg/heel vaak.

Fragment 1

	- -	-	+/-	+	++
1. Hoe verstaanbaar vond je deze spreker?	1	2	3	4	5
2. Hoe vaak heb je het accent van de spreker gehoord op tv/radio/internet?	1	2	3	4	5
3. Hoe vaak heb je het accent van de spreker gehoord buiten tv/radio/internet?	1	2	3	4	5
4. Uit welk land denk je dat deze spreker komt? Vul in:					

Fragment 2

	- -	-	+/-	+	++
5. Hoe verstaanbaar vond je deze spreker?	1	2	3	4	5
6. Hoe vaak heb je het accent van de spreker gehoord op tv/radio/internet?	1	2	3	4	5
7. Hoe vaak heb je het accent van de spreker gehoord buiten tv/radio/internet?	1	2	3	4	5
8. Uit welk land denk je dat deze spreker komt? Vul in:					

Fragment 3

	- -	-	+/-	+	++
9. Hoe verstaanbaar vond je deze spreker?	1	2	3	4	5
10. Hoe vaak heb je het accent van de spreker gehoord op tv/radio/internet?	1	2	3	4	5
11. Hoe vaak heb je het accent van de spreker gehoord buiten tv/radio/internet?	1	2	3	4	5
12. Uit welk land denk je dat deze spreker komt? Vul in:					

Appendix B

Transcript Listening Test

Fragment 1:

My hero isn't very famous, but she ought to be. She's Mary Anning, who was only 12 years old and from a poor family when she made an amazing discovery. She found the first dinosaur skeleton, that of an ichthyosaur, on the cliffs of Lyme Regis in the south of England. That was in 1811, and until then people had thought that it was impossible for an animal to become extinct. Because she was a woman and didn't have enough money for a proper education, she wasn't able to take part properly in the scientific community of the time. But she read as much scientific literature as she could and continued to search for fossils.

Fragment 2:

My hero is Kalash Satyarti, who has been campaigning against child slavery for years. He introduced a programme first called RugMark, now known as Goodweave, which puts tags on child-labour-free rugs made in factories. He has saved many thousands of children, over 80,000, from a terrible life of enforced labour in South Asia and helped them to get an education. He has often been physically attacked for helping children, for example for trying to free Nepalese children forced to work in a circus. He regularly risks his life to fight injustice; two of his colleagues have been murdered.

Fragment 3:

I'm really interested in nature and my hero, or heroine, is Rachel Carson because she first got people thinking about the way we humans are causing permanent damage to the Earth's ecosystems. She began as a biologist, specialising in writing about the sea, but she gradually became aware of the danger of using chemicals like DDT and the way they can harm the whole of the food chain, from the worm to a human! She wrote her classic book *Silent Spring* in 1962 to explain this to the general public, to explain how humans and nature depend upon each other.

Fragment 4:

My choice of hero isn't very original, I'm afraid, but he's the person I would most like to have met: John Lennon. But although I love his music, what I admire about him is his dedication to universal peace. It's amazing that the song *Imagine*, written in 1971, is still incredibly popular after all this time. It's about a world where everyone can be equal, a world with no wars, no divisions between countries, no greed, no hunger, no material possessions ... I'd like to meet him because he was a lifelong rebel, and although he could be a difficult person, he was original, clever and funny

Fragment 5:

Good morning. Today I'm going to talk about how to study. Now you probably think you know all about that, right? Right, so, what's the best way to study? Well, first of all, it's a good idea to have some kind of plan or timetable. This could be for the week or a longer revision timetable for an exam, from one month to six months. Yes, if you're studying for an important exam it's important to think long term. Draw up a timetable, but revise it often. If it's not going to plan, you may have to rethink it.

Fragment 6:

Next tip, think about your environment. Make sure the place where you are going to study is comfortable with enough light, air, etc. Not too hot, not too cold. Make sure there are no distracting noises around, such as television. If you think you concentrate better listening to music, experiment and see if it's really true. Some people really do seem to work better with music in the background, especially classical music, but for many people it spoils their concentration. However, if you have to work near a TV, you might have to use headphones to play music to drown out the sound of the TV.

Fragment 7:

Right, another important tip: planning breaks. Plan your study periods in chunks with regular breaks. Many people recommend half an hour of concentrated study, then a ten-minute break. But you can adjust this to suit you. If you study for too long at a stretch your mind will work less effectively, so be careful. In your short break, you can give yourself a treat, such as a cup of green tea. I wouldn't advise a chocolate bar as a treat – a sugar rush is not great for concentration. Eating lots of sweets has been proved to actually be counterproductive if you want to study.

Fragment 8:

OK, so next let's think about what you do when you're studying. Some people just read through their notes or textbooks and underline in pencil or highlight important bits. If this works for you, fine, but I'd suggest that it's better to write notes of some kind, so your mind is processing the information more. This way you are also producing material which will be useful for last-minute revision. I don't recommend super last-minute revision, but we'll talk about that later.

Fragment 9:

Listen to these examples of student notes. Which do you think are best? The first person has written important phrases at random, the next has a table with clear headings and boxes with notes, the next has summaries, and the last one has mind maps. In a mind map you've got the main topic in the circle, then lines coming off the circle attached to subcategories, then more lines to more details. Mind maps are great for showing the connections between different bits of information. They seem to work in the same way the brain works. So what do you think? To me, the first example wouldn't be very useful but the others are all fine – it depends on your personal preference.

Fragment 10:

Once again the organisation which promotes Australia, Tourism Australia, is advertising the best job in the world. The job of "outback adventurer" is for someone with a passion for outdoor life, and in the Northern Territory there are plenty of wide-open spaces. The job is for someone to find out the best adventures and jobs for young people on working holidays. You'd be getting close to wildlife, sleeping under the stars in a bush camp and flying over stunningly beautiful landscape in a hot air balloon. Your duties will include getting to know about aboriginal culture and eating traditional bushfoods, maybe including the famous witchetty insect larva.

Fragment 11:

Like the idea of 200 days of sunshine every year? Job number two is a park ranger in tropical Queensland. It's a wonderful state with ancient rainforests, the world's largest sand island and the

awesome Great Barrier Reef. Here your duties would include protecting and promoting native plants and animals, spectacular waterfalls, dinosaur fossils, untouched beaches and indigenous culture. You'd get paid to patrol the beaches of Lizard Island and sail around the beautiful coral reefs. You will learn how to scuba dive in shallow water and free dive just below the surface. You are going to live a life most people can only dream about.

Fragment 12:

Another island job is as 'wildlife caretaker' on Kangaroo Island in South Australia. If you love all kinds of animals, this is the job for you. The advert says you'll be able to talk to wallabies (a kind of small kangaroo), play with dolphins, cuddle koalas and sunbathe with seals on the unspoilt beach at Seal Bay. You would get about the island on foot, by bicycle, kayak or boat, taking photos and leaving only footprints. There is one potential difficulty, though. You'd need to be pretty brave since you might come face to face with great white sharks.

Fragment 13:

Maybe you are not quite so keen on the great outdoors and your talents are more journalistic. If you fancy feature writing, photography and making videos, you can apply for the position of lifestyle photojournalist for *Time Out* in Melbourne. You would be required to photograph and write about the city's coolest cafés and musical events. But you'd also cover tourist activities in the whole state of Victoria, including surfing on the Great Ocean Road, skiing at Mount Hotham or watching the little penguins at Phillip Island.

Fragment 14:

Are you a foodie? Do you know about food, as well as love eating it? If the answer's yes, you can apply for the role of 'taste master' in Western Australia. Your job would be to promote the best restaurants, pubs, wineries and breweries. You will get to meet the most famous chefs in Australia, tasting their fantastic recipes and you are free to experience some experimental cooking. You'd also catch fresh seafood off the beautiful coast and learn all about making wine and beer. A foodtastic job.

Fragment 15:

Finally, a fantastic job in Sydney. We've all seen those amazing firework displays in Sydney Harbour. Well, you could be one of the people making that happen next year. New South Wales is looking for a 'chief funster', who would be based in Sydney while travelling around the state and tweeting about the coolest things going on. This job would appeal to someone interested in everything: sports, the arts, entertainment, food. You'd also be involved in making the Sydney Festival, Mardi Gras and Vivid Festival as spectacular, and as fun, as possible.