RL Vol. 16.4 2018

Research in Language

WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

ŁÓDŹ 2018

EDITORS

Iwona Witczak-Plisiecka, Ewa Waniek-Klimczak, Jan Majer – University of Łódź
Faculty of Philology, Department of English Language and Applied Linguistics
90-236 Łódź, Pomorska 171/173

All papers published in "Research in Language" have been blind-reviewed by at least two independent readers

INITIATING EDITOR
Agnieszka Kałowska

EDITING AND TYPESETTING
Michał Kornacki

COVER DESIGN
Katarzyna Turkowska

Printed directly from camera-ready materials provided to the Łódź University Press

© Copyright by Authors, Łódź 2018
© Copyright for this edition by Uniwersytet Łódzki, Łódź 2018

Published by Łódź University Press

First Edition. W.09032.18.0.C

Printing sheets 8.875

ISSN 1731-7533
e-ISSN 2083-4616

Łódź University Press
90-131 Łódź, Lindleya 8
www.wydawnictwo.uni.lodz.pl
e-mail: ksiegarnia@uni.lodz.pl
phone (42) 665 58 63

CONTENTS

Léa Burin
Accommodation of L2 speech in a repetition task: exploring paralinguistic imitation 377

Rias van den Doel, Anne-France C. H. Pinget and Hugo Quené
Non-native attitudes to /θ/ and /ð/: A European case study 407

Yehudit Dror
Parenthetical clauses in the Qur'ān 429

Marta A. Nowacka
Back to orthoepia – spelling in pronunciation instruction: "Words Commonly Mispronounced" by learners of six L1s 451

Mariana Oleniak
Old English simile of equality: the highest degree of similarity 471

Ferran Suñer
The interplay of cross-linguistic differences and context in L2 comprehension 495

ACCOMMODATION OF L2 SPEECH IN A REPETITION TASK: EXPLORING PARALINGUISTIC IMITATION

LÉA BURIN
Paris Diderot University, France
lea.burin@univ-paris-diderot.fr

Abstract
Phonetic convergence is the process by which a speaker adapts his/her speech to sound more similar to his/her interlocutor. While most studies analysing this process have been conducted amongst speakers sharing the same language or variety, this experiment focuses on imitation between non-native and native speakers in a repetition task. The data is a fragment from the ANGLISH corpus designed by Anne Tortel (Tortel, 2008). 40 French speakers (10 male intermediate, 10 male advanced, 10 female intermediate and 10 female advanced learners) were asked to repeat a set of 20 sentences produced by British native speakers. Segmental (vowel quality), suprasegmental (vowel duration) and voice quality were analysed. Level of proficiency, gender and model talker were taken as independent variables.

Level appeared not to be a relevant parameter due to a high amount of inter-individual variability amongst groups. Somewhat contradictory results were observed for vowel duration and F1-F2 distance for male learners converged more than female learners. Our hypothesis that low vowels display a higher degree of imitation, and especially within the F1 dimension (Babel, 2012), was partially validated. Convergence in vowel duration in order to sound more native-like was also observed (Zając, 2013).

Regarding the analysis of voice quality, and more particularly of creaky voice, observations suggest that some advanced female learners creaked more than the native speakers and more in the reading task, which indicate, both linguistic idiosyncrasy and accommodation towards the native speakers. Low vowels seem also to be more likely to be produced with a creaky voice, especially at the end of prosodic constituents.

Keywords: imitation, convergence, vowel spectra, vowel duration, creaky voice

1. Introduction

1.1. Phonetic accommodation

Phonetic accommodation emerged from Communication Accommodation Theory (CAT) developed by Giles et al. in the early seventies. This theory argues that when people are involved in a conversation, they accommodate to others at different levels. One can mimic the postures, the mannerisms, or the facial expressions of interaction partners (Chartrand and Bargh, 1999: 893). Although people can imitate the persons they are interacting with spontaneously, it is mostly an unconscious process than can be strategic, and is motivated (Coupland and

Giles, 1988: 175) in order to, for instance, "improve communication at the social level and/or at the intelligibility level" (Garnier *et al*., 2013: 1).

The process of accommodation can be subdivided into three distinct sub-categories: divergence, maintenance and convergence.

1. One can diverge from one conversational partner in order to lower social approval (Giles et al., 1991: 32), and as a way of increasing social distance (Babel, 2012: 179). In this case, speakers will accentuate speech differences between themselves and their interlocutors.
2. One can maintain his/her natural way of speaking in order to maintain one's group identity (Bourhis, 1979).
3. One can adjust one's speech to sound more similar to his/her interlocutor (Babel, 2012: 177), leading to a decrease in the dissimilarities of acoustic-phonetic forms between talkers. This phenomenon has been the most frequently observed amongst conducted studies and is known as 'phonetic convergence'.

Speech rate and fundamental frequency have been proved to be the two features most prone to be imitated (Pardo, 2010; Sato *et al*., 2013). Many other prosodic features such as utterance length, pausal phenomena (Bilous and Krauss, 1998), vocal intensity (Black, 1949; Natale, 1975) or vowel duration (Zając, 2013) also appear to be subject to accommodation. The main communicative function that emerges from this concept of convergence is to lessen social distance, though accommodation can take place in all types of social situations. It can occur when speakers simply produce words (Goldinger, 1997, 1998; Goldinger and Azuma, 2004; Namy, Nygaard and Sauerteig, 2002), but is mostly observable in socially rich, dyadic conversations (Natale, 1975; Pardo, 2006; Pardo *et al*., 2010, Babel, 2011).

Welkowitz *et al*. (1972) showed that pair talkers who perceived themselves as being similar converged more towards each other in terms of vocal intensity than randomly-paired talkers. Giles (1973) analysed a conversation between an inspector and a traveller in a train and observed that the latter would converge towards the speech of the former. Consequently, the way people speak can lead to determine which person is socially dominant in a conversation (Nilsenová and Swerts, 2012: 77). It has been found that, in dyadic conversations, each person is assigned a role and people are more likely to converge towards the person who has more 'power' (Watzlawick *et al*. 1967).

Gender has also been found to play a role in the process of accommodation. Women appear to converge more than men (Eisikovits, 1987; Namy *et al*., 2002; Babel, 2012; Babel *et al*., 2014). A possible explanation for this pattern is that "women might be more sensitive to indexical features[1] of talkers, which could lead to greater convergence" (Pardo, 2006: 2388). In addition, speakers, whether male

[1] Indexical features are data about a person such as physical, social, and psychological characteristics, e.g. age, gender, social status, and emotional state (Abercrombie, 1967; Laver, 1989; Laver and Trudgill, 1979).

or female, have a tendency to converge more towards male interlocutors (Bilous and Krauss, 1988; Gallois and Callan, 1988; Willemyns *et al*. 1997).

Accommodation, and mostly convergence, may also be affected by the talker's perceived attractiveness. People tend to converge towards an interlocutor they appreciate and by whom they want to be appreciated (Giles *et al*., 1991, Byrne, 1997; Chartrand and Bargh, 1999; Babel, 2009; Babel, 2012). Imitating others has, thus, a clear social function in that it reflects the degree of affiliation we feel and want to elicit in other people (Lakin *et al.* 2003).

In the literature on second language acquisition (SLA), Giles and Johnson (1987) have found evidence that a non-native speaker will be likely to imitate a native speaker if they both share significant social identities, related to ethnicity or not. However, Zuengler (1982) demonstrated that the L2 pronunciation can vary, by diverging or converging, if a native-English-speaking interlocutor conveys negative or positive attitude towards the ethnic group the non-native speaker belongs to. In the investigation he conducted, non-native speakers who perceived threat would phonetically diverge if they firmly identified as ethnic group members, or if they wanted to defend their ethnic solidarity. Zając (2013) conducted an experiment to determine how phonetic imitation can, or cannot, be influenced by the model talker being a native or a non-native speaker of English. She found that informants converged towards the native model talker and diverged from the non-native model talker. As an explanation, Zając suggested that subjects might have been aware of the foreign accent of the nonnative speaker and have diverged from her in order to distance themselves from other foreign-accented talkers, and converged towards the native model in order to sound more native-like. This provides strong evidence that non-native speakers use identical strategies to native speakers in order to lower social distance.

1.2. The current experiment

This study investigates deliberate phonetic accommodation between native and non-native speakers. The data is a fragment from the ANGLISH corpus designed by Anne Tortel (Tortel, 2008). 40 French learners of English (10 male intermediate, 10 male advanced, 10 female intermediate and 10 female advanced learners) were asked to repeat and imitate a set of 20 sentences produced by two British native speakers (one male and one female). Convergence with the realisations of read speech recorded before the repetition task are analysed.

Previous findings on accommodation in spontaneous speech indicate that female informants tend to converge more than male informants (Namy *et al*., 2002; Babel, 2009; Babel *et al*., 2014), and advanced learners to converge more than intermediate learners for "greater L2 usage and proficiency are associated [...] with increased L2 production experience" (Best and Tyler, 2007: 20). We expect female speakers to converge more, whatever their level. On the contrary we expect level to play a role in the type of accommodation observed. Convergence in vowel duration in order to sound more native-like (Zając, 2013)

and accommodation for low vowels, especially within the F1 dimension (Babel, 2009, 2012), are more likely to be observed. Euclidean distance for normalized F1-F2 will be compared for level of proficiency and gender.

Our last section will discuss specific features traditionally ignored or rejected outside the perimeter of phonetic features and called 'paralinguistic' features, such as creaky voice. Indeed, this sociolinguistic feature consistently appears for some French advanced female learners. Preliminary observations indicate that some advanced female learners creak more than the native speakers, and more in the reading task, which could indicate both linguistic idiosyncrasy and accommodation towards the native speakers. Low vowels also seem to be more likely to be produced with a creaky voice, especially at the end of prosodic constituents.

2. Methodology

2.1. Corpus

The data used for this research study was collected from the ANGLISH corpus (Tortel, 2008). 20 French intermediate (10 female, 10 male) and 20 advanced learners (10 female, 10 male) of English[2] were asked to perform two different tasks: first the reading of 4 passages, each related semantically, then the repetition of these same sentences after a native model (Tortel, 2013: 234).

For the sake of naturalness, the repetition task was only carried out once. This has an unfortunate consequence, as the influence of gender cannot be compared for each subject. To make up for this, two model voices have been used for the native realisations: one British female speaker and one British male speaker. The distribution of the corpus between the male and the female voice is, however, slightly unbalanced. The first passage was read by a female native British speaker and the three others were read by the same male native British speaker.

The data used comprised the reading and repetitions of these extracts, along with the production of the two native model talkers, resulting in 815 recordings lasting in average 4 seconds. Subjects could read the text corresponding to the soundfile that was played during the repetition task.

All recordings were made in an anechoic chamber at a frequency of 44KHz and 16-bit resolution at the *Parole et Langage* laboratory in Aix-en-Provence, and were done using a headset microphone.

[2] Advanced learners were students majoring in English at university, and intermediate learners were adult speakers who did not specialise in English.

2.2. Extraction of the acoustic components

Formant values and vowel duration of all monophthongs occurring in the corpus were extracted, resulting in the analysis of 12,610 phonemes.

Front vowels /ɪ, i, iː, e, æ/, back vowels /uː, ʊ, ʌ, ɔː, ɒ, ɑː/ and the central vowel /ɜː/ were analysed. Vowels are coded through the whole study according to Wells' Lexical Set (Wells, 1982).

F1 and F2 values were extracted using a Praat script for each target vowel at different percentages of the distance from the onset to the offset (respectively 25%, 50% and 75%) of the vowels. The average value was then calculated to obtain a single, mean value for each vowel.

Duration was extracted for each target vowel and for each speaker with a Praat script designed by Mietta Lennes.[3]

2.3. Analysing the acoustic components

The point of this research study is to determine how the participants' production changed after auditory exposure to the model talkers. The method used is the same as the one used by Babel (2012) which consists in calculating the distance value of how the production of a particular vowel evolves through the course of the experiment. This distance value is referred to as the 'difference in distance'.[4]

The difference in distance was first compared between level, then between gender, and finally between gender and level at the same time. The variable model talker was kept into account, as the difference in distance may not display the same results depending on which model talker learners were repeating.

All vowels were normalized using the Lobanov method.

2. Results

3.1. Difference in distance

The majority of subjects fall on the negative side of the scale (see Figure 1), indicating a general tendency to converge towards the vowel spectra of the model talkers. Results from a series of *t*-tests showed that the overall difference in distance values between participants were, however, not significant [M = -0.0264, t = -1.422, $p = 0.078$]. It appeared to be also the case for the majority of the vowels. Only three vowels out of twelve displayed significant difference in distance values: the DRESS [M = -0.155, $p < 0.001$], START [M = -0.247, $p < 0.01$] and FORCE [M = -0.132, $p < 0.01$] vowels. These observations demonstrate that there were no effects of imitation across the entire dataset and for each vowel. The

[3] SpeCT – The Speech Corpus Toolkit for Praat – http://www.helsinki.fi/~lennes/praat-scripts/

[4] See Babel (2012) for a detailed explanation of how the difference in distance value is calculated.

following analyses explore whether the observed pattern appeared to be the same across groups.

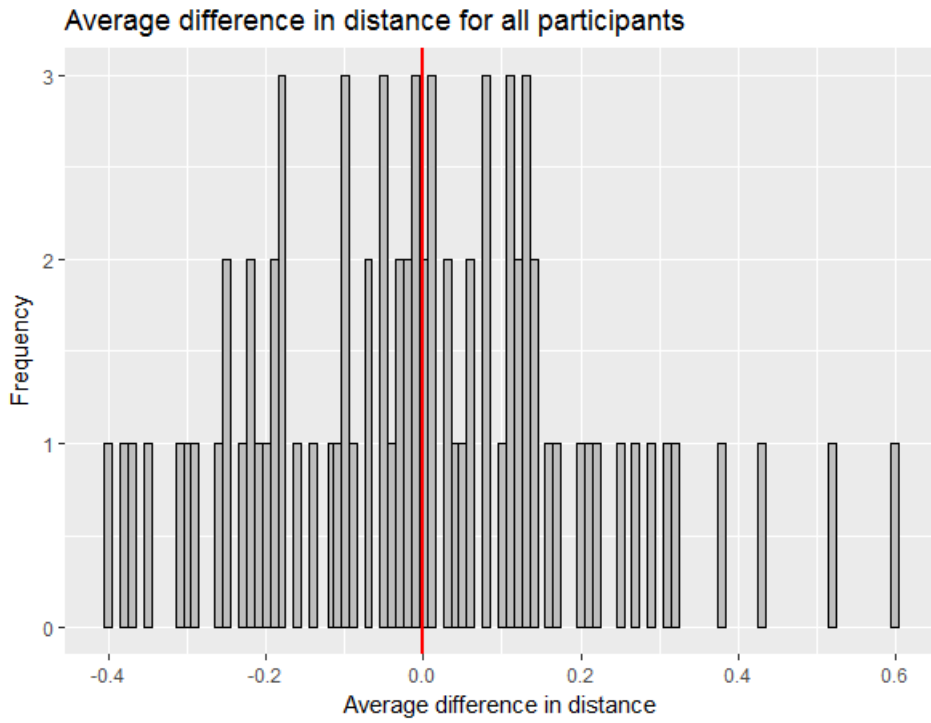

Figure 1. Histogram of all participants' averaged difference in distance value. The red vertical line indicates the 0 point of no change; a positive value indicates vocalic divergence, while a negative value indicates vocalic convergence. The number of subjects is represented on the vertical axis

3.1.1. Across level

The majority of advanced learners fall on the negative side on the scale (see Figure 2), indicating a tendency to converge towards both model talkers. In addition, half of the intermediate learners fall on the negative side of the scale and the other half on the positive side. Half of the intermediate learners have therefore diverged from the model talkers, while half have converged towards them. The difference in distance values across level were, however, not significant [M(Adv) = -0.024, M(Int) = -0.029, p = 0.55]. The hypothesis that advanced learners accommodate more than intermediate learners cannot be validated.

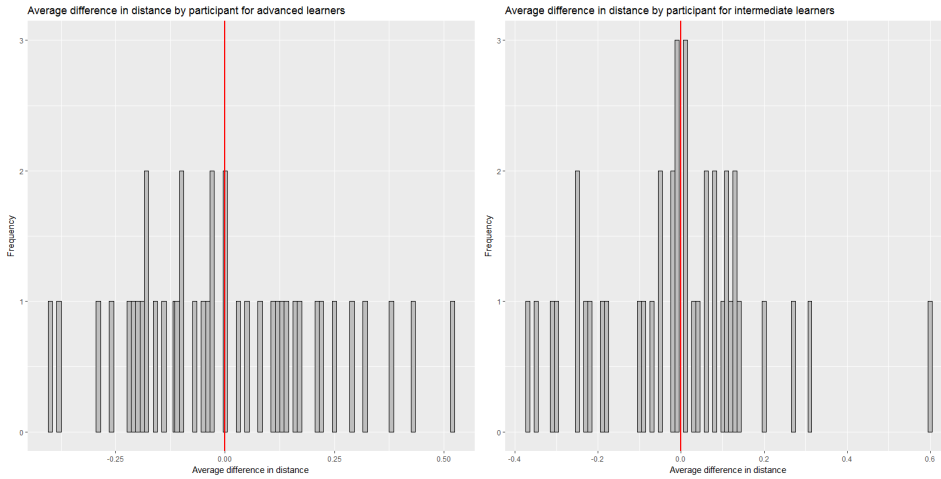

Figure 2. Histograms of each advanced and intermediate learner's averaged difference in distance value. Advanced learners are on the left panel and intermediate learners are on the right panel. The red vertical line indicates the 0 point of no change; a positive value indicates vocalic divergence, while a negative value indicates vocalic convergence

3.1.2. Across gender

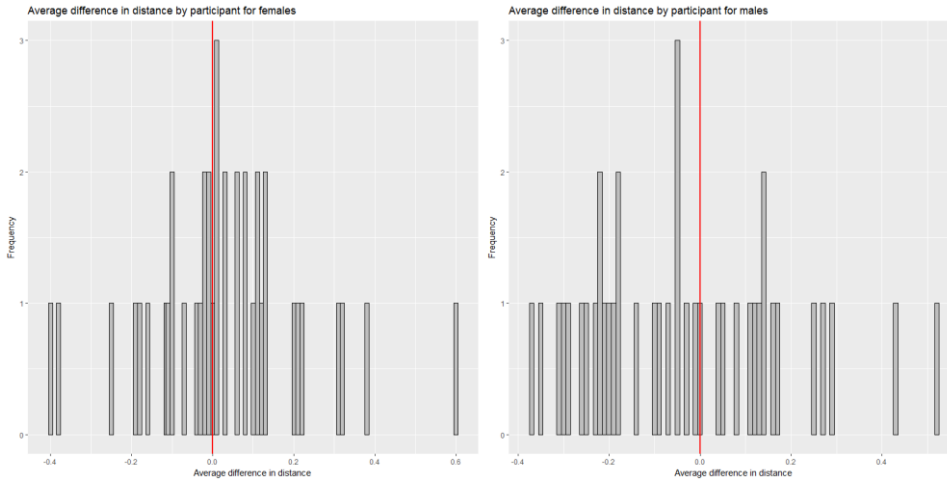

Figure 3. Histograms of each male and female learner's averaged difference in distance value. Female learners are on the left panel, and male learners are on the right panel. The red vertical line indicates the 0 point of no change; a positive value indicates vocalic divergence, while a negative value indicates vocalic convergence

The majority of female participants fall on the positive side of the scale (see Figure 3), indicating divergence from the model talkers. Conversely, the majority of male participants fall on the negative side of the scale, indicating a tendency to converge

towards the vowel spectra of the model talkers. Statistical analyses demonstrated, however, that the difference in distance values across gender was not significant [M(Female) = -0.008, M(Male) = -0.045, p = 0.84]. The hypothesis that female speakers tend to accommodate more than male speakers cannot be validated.

3.1.2. A cross-analysis between level, gender and model talker

The following analyses explore the differences in imitation and the different patterns across groups that arose through the experiment. The results of an ANOVA on the full design are first reported. Subsequently, this is broken down into cases when participants were repeating the male or the female model talker.

The full design of the experiment was a 2 (Model: Female or Male) × 2 (Gender: male or female) × 2 (Level: advanced or intermediate) × 12 (Vowel: KIT, DRESS, TRAP, LOT, STRUT, FOOT, FLEECE, GOOSE, NURSE, START, FORCE, HAPPY) factorial design. There were main effects of Vowel [F = 4.004, p < 0.001]. The two-way Vowel × Model [F = 10.641, p < 0.01] and Vowel × Gender [F = 4.333, p < 0.001] interactions appeared to be significant as well, such as the three-way Vowel × Gender × Level [F = 1.917, p < 0.05] and Vowel × Gender × Model [F = 4.343, p < 0.001] interactions. The four-way Vowel × Gender × Level × Model interaction was proved to be insignificant [F = 0.893, p = 0.53]. A post-hoc *t*-test comparing the difference in distance values between male [M = -0.064] and female model talker [M = 0.019] revealed significantly more imitation in the task of repeating the male model talker [t = -2.197, p < 0.05], which corroborates previous findings (Bilous and Krauss, 1988; Gallois and Callan, 1988; Willemyns *et al*. 1997).

The following analyses investigate the two three-way interactions that appeared to be significant (see Figure 4 and Table 1).

Depending on the type of interactions, either H1 or H2 is validated but not both at the same time. Moreover, if we compare the number of vowels imitated between advanced and intermediate learners, it appears that advanced learners imitated more vowels than intermediate learners, which correlates with Best & Tyler's (2007:20) observation that "greater L2 usage and proficiency are associated […] with increased production experience". Another interesting observation is that the learners' tokens for FLEECE do not display imitation in the majority of cases. This might be due to the fact that KIT and FLEECE are assimilated to one L1 sound, a French-like [i][5]. As predicted by the Perceptual Assimilation Model – PAM (Best, 1995), discrimination is thus very poor. Production and the ability to accommodate are found to be affected.

[5] The tense allophonic realisations for happY are not discussed here.

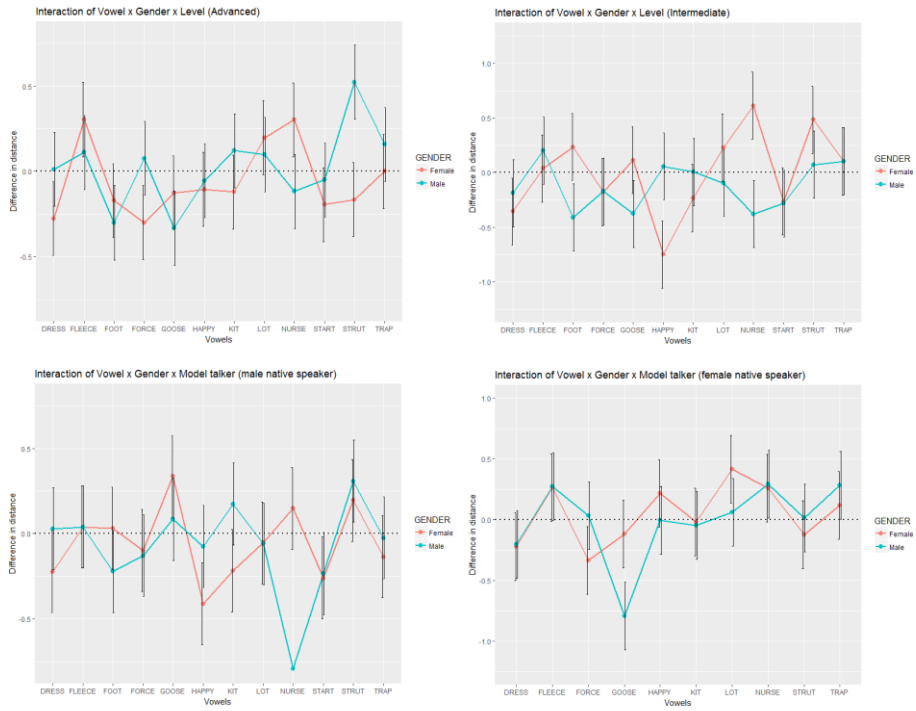

Figure 4. Graphs representing the four types of interaction. The Vowel x Gender x Level (Advanced) interaction is displayed on the top left panel; the Vowel x Gender x Level (Intermediate) interaction is displayed on the top right panel; the Vowel x Gender x Model (Male native speaker) is displayed on the bottom left panel, and the Vowel x Gender x Model (Female native speaker) is displayed on the bottom right panel. The difference in distance measure on the y-axis indicates the amount of phonetic imitation. A value of zero shows no change in vowel production as a result of auditory exposure to the model talker; a positive value indicates vocalic divergence, and a negative value demonstrates vocalic convergence. Error bars are calculated based on the standard deviation of the difference in distance measurements.

Table 1. Summary of the results of the cross-analyses. 1 and 2 indicate that the male learners were more likely to imitate the NURSE vowel than the female learners (posthoc Tukey HSD = ($p < 0.05$)). 3 indicates that the male learners were more likely to imitate the GOOSE vowel than the female learners (posthoc Tukey HSD = ($p < 0.01$))

Hypothesis **Interaction**	**H1** ***(female converge more than male)***	**H2** ***(low vowels more imitated than high vowels)***
Vowel, Gender, Level (advanced)	✓	✗
Vowel, Gender, Level (intermediate)	✗[1]	✓
Vowel, Gender, Model (male)	✗[2]	✓
Vowel, Gender, Model (female)	✓[3]	✗

3.2. Normalization

Analyses so far depict the general nature of vowel imitation based on the difference in distance measurements. Normalized formant plots were created to observe the direction of imitation. The three-way interaction vowel by gender by level was found to be insignificant, implying that level is not a meaningful factor. Therefore, it was excluded from the following analyses. This analysis allows for the comparison of the average realisation of learners (in the repetition and reading tasks) compared to that of the models (see Figure 5). It should be borne in my mind that the average represented for each learner actually corresponds to a greater variability (20 female and 20 male learners).

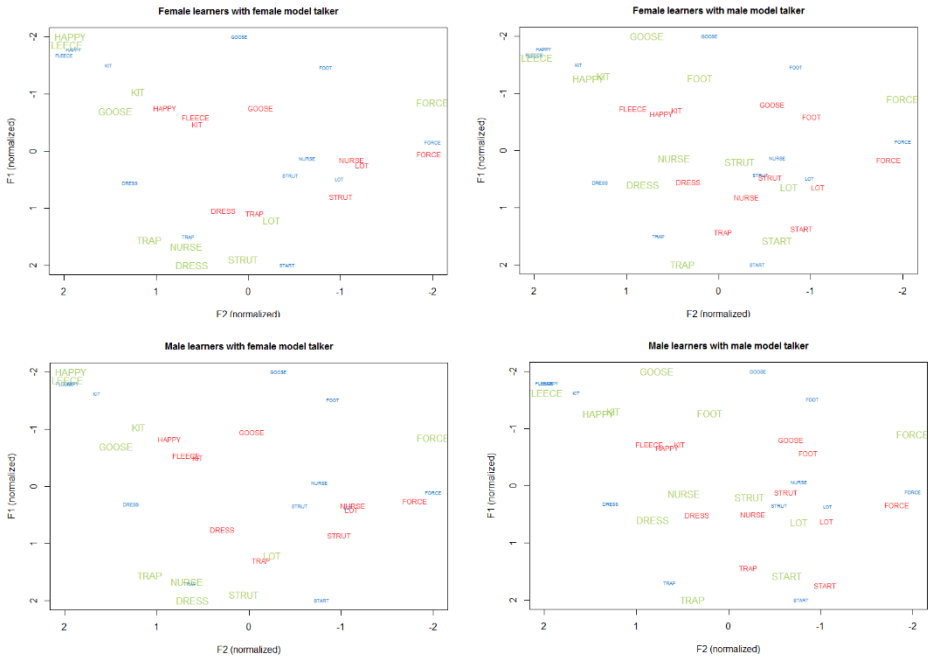

Figure 5. Formant plots displaying the direction of spontaneous phonetic imitation across gender. Female speakers are on the top panels and male learners on the bottom panels. The average baseline of the model talkers' productions is displayed in the green font. The average baseline of participants' productions in the read task is plotted in the blue font and their production in the repetition task is plotted in the red font.

The main observation emerging from the analysis of normalized formant plots is that, in all cases, the production of the FLEECE and HAPPY vowels is more similar to that of natives (green) in the reading task (blue) than in the repetition task (red). In the latter, the three realisations for HAPPY, KIT and FLEECE cluster, suggesting that the distinction is not perceived. From a psycholinguistic

point of view, general findings show that L2 learners tend to have difficulty differentiating two L2 sounds that do not contrast in the L1, especially when they are phonetically similar to an L1 sound. Several models have been proposed to explain this difficulty in perceiving L2 sounds, such as the Perceptual Assimilation Model – PAM (Best, 1995), the Speech Learning Model – SLM (Flege, 1995) or the Native Language Magnet Theory – NLM (Kuhl and Iverson, 1995). Essentially, it appears that the degree to which L2 contrasts are difficult to perceive depends on how relevant sounds are distributed in the L1 and the L2. How learners perceive L2 sounds can, thus, affect the production of these same sounds. However, this observation indicates that learners are more proficient at marking this distinction in the reading task rather than in the repetition task.

Furthermore, when low vowels appeared to be imitated, they would also display a higher degree of imitation in the F1 dimension, corroborating Babel's (2009) results.

3.3. Vowel duration

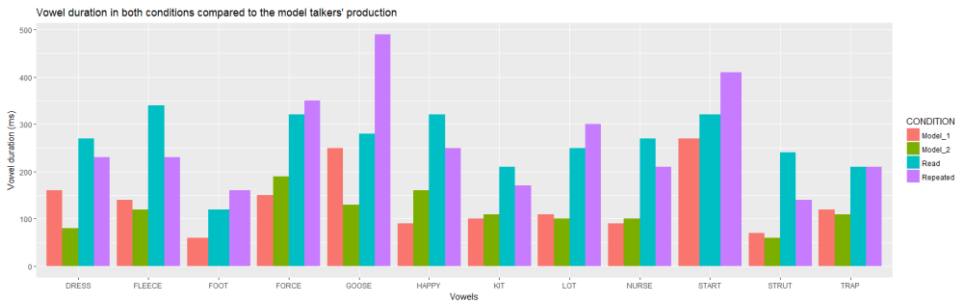

Figure 6. Histogram displaying the overall duration of each vowel in both conditions compared to the model talkers' production. The male model talker's vowel durations are displayed in red, while the female model talker's vowel durations are displayed in green. Participants' vowel durations in the read condition are displayed in blue and those in the repeated condition are displayed in purple.

The non-native speakers produce longer vowels than the native speakers, whatever the task (see Figure 6). Native speakers always speak at a faster rate than non-native speakers and, since speaking at a slower rate results in the production of longer vowels, longer vowels for learners were predictable. We can observe that, overall, the male native speaker [M = 77. 403] produces longer vowels than the female native speaker [M= 74.23]. A *t*-test determined that the difference between the two model talkers was, however, not significant [p = 0.60]. Results of another *t*-test showed that the difference regarding the learners' vowel duration between the two tasks was insignificant [p = 0.88].

That learners increased the duration for some of the long vowels is interesting, as it might be due to the fact that they have perceived the dichotomy that exists between long and short vowels in English. This argument only applies for

intermediate learners though, for advanced learners have received phonological training and are undoubtedly aware of this distinction. Yet it appears not necessarily to result in better production.

The following analyses explore the differences in imitation and the patterns that arose across groups. First, we have tested whether vowel duration was influenced by Vowel × Gender × Level. The independent variable model talker will be taken into account in a further analysis. Results from ANOVA indicate main effects of Vowel [F = 377.175, $p < 0.001$], Gender [F = 182.674, $p < 0.001$] and Level [F = 127.931, $p < 0.001$]. These factors were also significant as two-way Vowel × Gender [F = 3.157, $p < 0.001$] and Vowel × Level [F = 6.606, $p < 0.001$] interactions. A three-way Vowel × Gender × Level [F = 1.634, $p = 0.08$] interaction was not found. The data was then subset into gender and level and submitted to separate statistical analyses. All results of the statistical analyses are reported in the appendix section.

3.3.1. Level

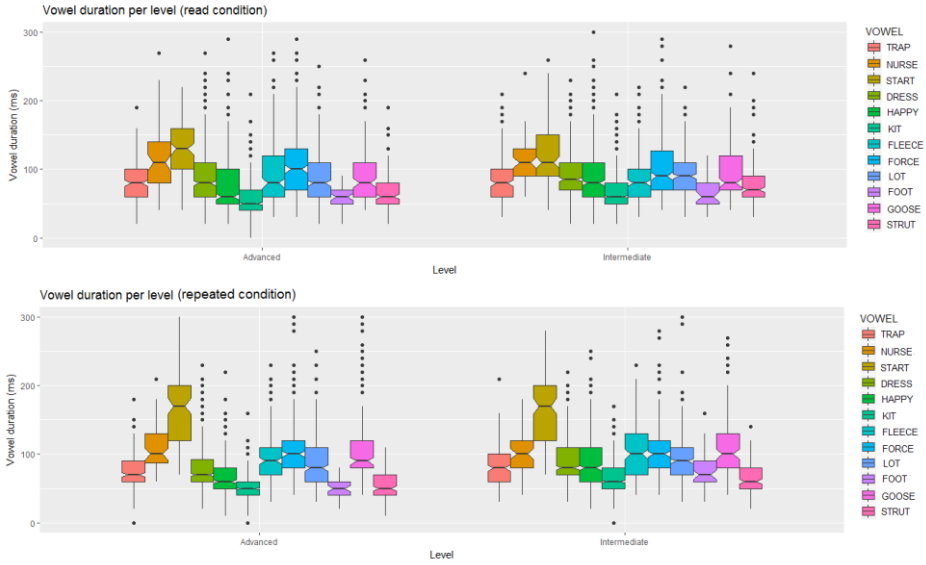

Figure 7. Boxplots displaying vowel duration per level in both tasks

In the reading task, advanced learners tend to produce shorter short vowels and longer long vowels than intermediate learners (see Figure 7). This indicates that advanced learners are more aware of the distinction that exists between short and long vowels in English and that they are, thus, more proficient in the production of durational characteristics. However, statistical analyses (see Table 2) demonstrated that long vowels were not significantly longer when produced by the advanced learners. Most of the short vowels appeared, however, to be

significantly shorter. This was not the case for the GOOSE vowel. The DRESS and LOT vowels appeared to have similar duration across groups. These observations indicate that advanced learners produce shorter short vowels than intermediate learners, indicating that their production is more similar to those of the model talkers. No firm conclusion can be drawn regarding the production of long vowels.

In the repetition task, it appears that advanced learners produce shorter vowels than intermediate learners, except for the NURSE and START vowels. Results of the statistical analyses (see Table 3) demonstrate that the majority of the short vowels were significantly shorter in that case, except for the FLEECE, FORCE and GOOSE vowels. The same conclusion drawn from the observation of vowel duration in the reading task is also valid in this case.

If we now compare the advanced learners' production between the two tasks we can observe a tendency to vowel shortening. Almost half of the short vowels were found to be significantly shorter (see Table 4 in the appendix). Conversely, the START and GOOSE vowels were proved to be longer in the repetition task. Conclusions that can be drawn from these observations are that all vowels which appear to be significantly shorter between the two tasks are short vowels. Vowels which are significantly longer, or which are not proved to be shorter in the repetition task, are long vowels.

Regarding the intermediate learners' production, we can observe a higher amount of vowel lengthening in the repetition task. Results of the statistical analyses displayed in Statistical analyses (see Table 5) demonstrate that the START, FLEECE and GOOSE vowels are, indeed, significantly longer in this task. This is not the case for the FOOT and FORCE vowels. The STRUT and NURSE vowels appear to be significantly shorter. The TRAP, DRESS, HAPPY, KIT and LOT vowels do not seem to vary in terms of duration between the two tasks. Statistical analyses indicate, however, that the KIT vowel is significantly shorter [M(read) = 68.576, M(rep) = 65.117), $p < 0.001$]. The other vowels do not display any significant amount of shortening or lengthening. Most of the long vowels are lengthened in the repetition task, which could mean that intermediate learners have perceived the dichotomy between short and long vowels. We can also observe that the durational characteristics of short vowels are more subject to imitation than those of long vowels.

Overall, advanced learners tend to produce shorter short vowels. Speaking at a faster rate results in the production of shorter vowels, and more proficient learners tend to speak faster than less proficient learners. This observation confirms that "greater L2 usage and proficiency are associated […] with increased production experience" (Best and Tyler, 2007: 20). In addition, they shorten the majority of short vowels whereas intermediate learners have not. Since both model talkers produced shorter vowels than learners, this observation suggests that advanced learners converge towards the native speakers to a higher degree than intermediate learners.

Vowels displaying lengthening in the repetition task across levels are all long vowels. If convergence had occurred, the duration of long vowels would have been shortened. However, these results indicate that all subjects perceived and noticed the distinction that exists between long and short vowels in English. We would suggest that, regarding the production of long vowels, perceptual accommodation occurred, but not phonetic accommodation.

3.3.2. Gender

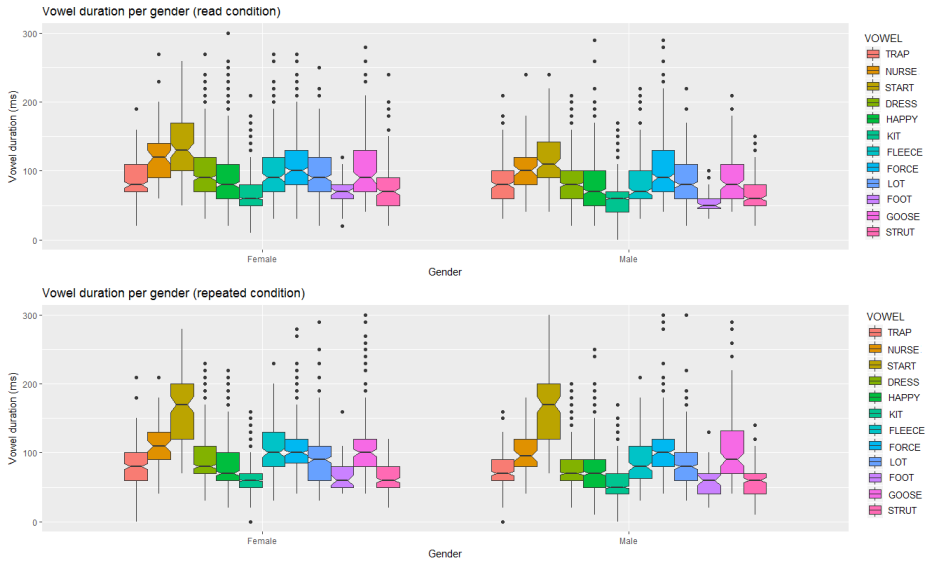

Figure 8. Boxplots displaying vowel duration per gender in both tasks.

Female participants tend to produce longer vowels in the reading task, except for the FORCE vowel (see Figure 8). These observations were proved to be true, since only the FORCE vowel appeared not to be significantly shorter (see Table 6). Moreover, the findings indicate the tendency for female participants to produce longer vowels than male participants.

As in the reading task, it seems that female participants produce longer vowels than male participants in the repetition task. All vowels appear to be longer except the START vowel which seems to have the same duration across groups. Results of a series of *t*-tests (see Table 7) demonstrate that all vowels are significantly longer when produced by the female participants, except for the GOOSE and FORCE vowels. The same conclusion drawn from the analysis of vowel duration in the reading task appears to be valid.

If we now compare the female participants' production between the two tasks, we can observe that the FLEECE, FORCE, GOOSE and START vowels were subject to lengthening and other vowels to shortening. However, statistical

analyses (see Table 8) demonstrated that only the TRAP, DRESS, HAPPY, KIT and STRUT vowels were significantly shortened. In addition, only the FLEECE vowel was proved to be longer. These observations indicate that female participants shortened the duration of most of the short vowels, and, thus, partly converged towards both model talkers. No definitive conclusion can be drawn from the analysis of long vowels.

Regarding the male participants' production between the two tasks, it appears that, as for female participants, the FLEECE, FORCE, GOOSE and START vowels have been subject to lengthening, in addition to the FOOT vowel. Other vowels seem to display shortening between the two tasks. These observations were statistically proved (see Table 9), except for the NURSE, LOT, FLEECE, FORCE and FOOT vowels. As for female participants, male participants shortened the duration of most of the short vowels. Accommodation towards both model talkers was, only partially manifested. No final conclusion can be drawn from the analysis of long vowels.

Overall, the analysis of vowel duration across gender indicates that female participants produce longer vowels than male participants. All participants shortened the duration of most of the short vowels in the repetition task, indicating that they converged towards both model talkers. No categorical conclusion can be drawn from the analysis of long vowels despite the fact that they seem to have been lengthened but, for most of them, it was not proven statistically. Nonetheless, this indicates that all participants have perceived and noticed the distinction between long and short vowels. We can conclude that short vowels are more subject to imitation than long vowels.

3.3.3. A cross-analysis between, level, gender and model talker

Previous observations showed that the male model talker tends to produce longer vowels than the female model talker. Vowels displaying longer duration are the DRESS, FLEECE, GOOSE, LOT, STRUT and TRAP vowels. No comparison is possible for the FOOT and START vowels for they are not produced by the female model talker.

The data was subcategorized to observe whether participants across level and gender would produce longer DRESS, FLEECE, GOOSE, LOT, STRUT and TRAP vowels when repeating the male model talker.

Advanced learners produced longer vowels when repeating the male native speaker, except for the FORCE, HAPPY, KIT, NURSE and STRUT vowels (see Figure 9). However, we have seen that the male model talker tended to produce shorter FORCE, HAPPY and KIT vowels. This is evidence that advanced learners have converged towards both model talkers, except for the NURSE and STRUT vowels. These observations were statistically proved (see Table 10), except for the NURSE, KIT and STRUT vowels. We can conclude that, overall, advanced learners have imitated both model talkers in terms of vowel duration.

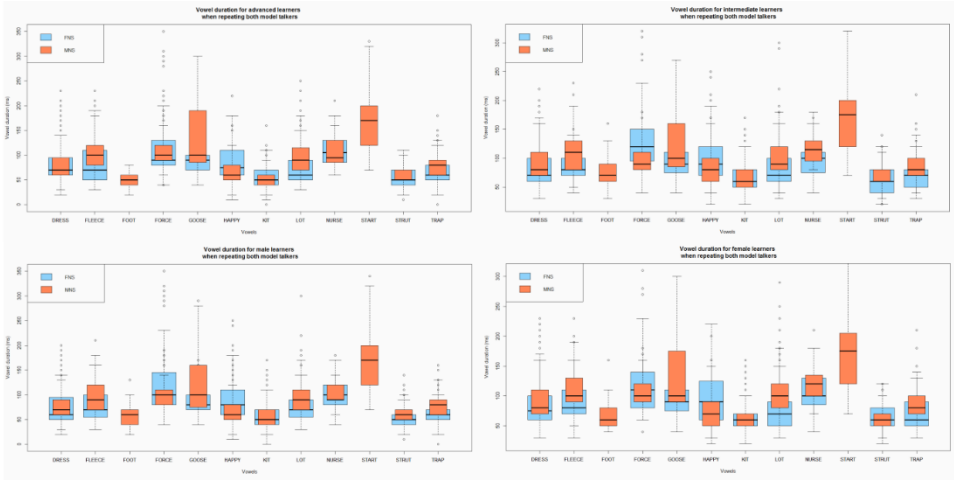

Figure 9. Boxplots displaying vowel duration when repeating both model talkers. The advanced learners' production is displayed on the top left panel; the intermediate learners' production is displayed on the top right panel; the male learners' production is displayed on the bottom left panel, and the female learners' production is displayed on the bottom right panel. FNS corresponds to the female native speaker while MNS corresponds to the male native speaker. Boxes in orange indicate vowel duration in the task of repeating the male native speaker and those in blue indicate vowel duration in the task of repeating the female native speaker. The FOOT and START vowels are not displayed in blue for they were not produced by the female native speaker.

Regarding the intermediate learners' production, it appears that they produced longer vowels when repeating the male model talker, except for the FORCE and HAPPY vowels. The KIT vowel displays no change in duration between the two tasks. Statistical analyses (see Table 11) verified these observations: only the STRUT vowel was not found to be significantly longer. Hence, intermediate learners have imitated both model talkers.

Female learners appear to produce longer vowels in the task of repeating the male native speaker, except for the FORCE, HAPPY and STRUT vowels. The KIT vowel displays no change in duration between the two tasks. These observations were statistically proved (see Table 12) except for the NURSE and STRUT vowels. We can conclude that female learners have, overall, imitated both model talkers.

Male learners seem to produce longer vowels in the task of repeating the male model talker, except for the FORCE and HAPPY vowels. The KIT vowel seems not to display any change in duration between the two tasks. A series of *t*-tests (see Table 13) concurred with these observations, except for the NURSE and STRUT vowels. Male participants have thus, overall, imitated both model talkers.

Among the conclusions that can be drawn from the analysis of vowel duration, we find that all groups have, overall, imitated both model talkers. Furthermore, when the male model talker produces longer vowels participants would, in turn, produce longer vowels. An interesting observation is that the KIT vowel usually

displays no change in duration between the two tasks. This might be due to the fact that this L2 sound has been perceived according to its articulatory similarities with French /i/, or that, since participants could read the text corresponding to the soundfile that was played, there might have been some grapho-phonemic influence. Learners then produced the KIT vowel in the same way as the HAPPY vowel in terms of duration. We can conclude that there has been a transfer effect from the L1 to the L2 regarding the production of this vowel.

To determine whether participants converged more towards the male or the female model talker, mean duration for all vowels was calculated for all participants in the two repetition tasks, and was then subtracted from the mean duration of the model talkers' production. A lower difference in vowel duration value indicates that the participants' vowel duration is closer to the model talkers' production. Results show that the difference value is lower in the task of repeating the female model talker (see Table 14). Participants' production in terms of vowel duration was, then, closer to the production of the female model talker, meaning that they converged more towards her than towards the male model talker.

Table 14. Difference in duration values between the two repetition tasks. The difference in duration is expressed in millisecond.

Model talker	**Difference value (ms)**
Male	6.599
Female	3.655

The same analysis was replicated for each group separately to determine whether the general tendency to accommodate more towards the female model talker was also valid across groups.

Table 15. Difference in duration values across groups in the two repetition tasks. The difference in duration is expressed in millisecond.

	Difference value (ms)			
Model talker	Advanced learners	Intermediate learners	Female learners	Male learners
Male	2.895	10.303	11.414	1.784
Female	1.057	6.252	6.843	0.467

Results indicate that advanced learners converged more towards the male model talker but all other groups seem to have converged more towards the female model talker because the difference in duration values appear to be lower (see Table 15). This confirms that, overall, participants converged more towards the female model talker. This table also suggests that advanced learners converged more than intermediate learners, and male participants more than female participants.

The same analysis was run through for each vowel separately to determine which vowels, or vowel category, appear to be the most subject to deliberate accommodation. Ideally, a value of 0 would indicate that the participants' mean vowel duration is equal to the model talkers' vowel duration. Negative values indicate that participants' mean vowel duration is shorter than the model talkers' production. When the difference in vowel duration is highly distant from 0, it means that the participants' production is highly distant to the model talkers' production. Comparison for the FOOT and START vowels is not possible since these vowels were not produced by the female model talker.

Table 16. Difference in duration values for all vowels across groups in the two repetition tasks. The difference in duration is expressed in millisecond.

		Difference value (ms)		
Model talker	**Vowels**	Advanced learners	Intermediate learners	Female learners
Male	DRESS	2.9	14.642	15.525
	FLEECE	-7.7	1.2	4.6
	FOOT	3.25	23.25	18.25
	FORCE	0.3	-8.6	-1.575
	GOOSE	15.5	3.5	17.75
	HAPPY	2.127	20	13.909
	KIT	4.97	15.96	12.608
	LOT	15.5	22.167	23.889
	NURSE	13.5	28.5	30
	START	-12.25	-13.25	-11.33
	STRUT	1.9	13.8	10.6
	TRAP	-5.25	2.467	2.75
Female	DRESS	12	15.625	18.75
	FLEECE	-6.333	-4.917	1.333
	FORCE	3	15.891	10
	GOOSE	-22.75	-17.5	-19.25
	HAPPY	-10.2	6	-0.4
	KIT	-4.642	4.309	1.825
	LOT	6.25	12.283	10.667
	NURSE	31	14.5	27.25
	STRUT	6.125	13.5	14.625
	TRAP	-3.875	2.833	3.625

Advanced learners appear to accommodate more towards the male model talker, as seen previously (see Table 16). Difference in duration values are lower for the DRESS, FORCE, GOOSE, HAPPY, NURSE and STRUT vowels when repeating the male model talker, and lower for the FLEECE, KIT, LOT and TRAP vowels when repeating the female model talker. Advanced learners converged more

towards the male model talker regarding the production of these vowels. In the task of repeating the male model talker, the FORCE vowel displays the lowest difference in duration value, meaning that this vowel is the one which is the most subject to accommodation. This is also the case in the task of repeating the female model talker.

Intermediate learners converged equally towards both model talkers, as 5 out of 10 vowels display a lower difference in duration value in both cases. The DRESS, FLEECE, FORCE, GOOSE and TRAP vowels show a lower difference in duration values when repeating the male model talker, meaning that participants converged more towards him in terms of vowel duration for these vowels. Intermediate learners converged more towards the female model talker regarding the HAPPY, KIT, LOT, NURSE and STRUT vowels. In addition, the vowel displaying the highest degree of convergence in the task of repeating the male model talker is the FLEECE one, while in contrast it is the TRAP vowel that displays the greatest degree of convergence in the task of repeating the female model talker.

Female participants appear to have converged more towards the female model talker since 6 (FLEECE, HAPPY, KIT, LOT, NURSE and TRAP) out of 10 vowels display a lower difference in duration values. Conversely, the DRESS, FORCE, GOOSE and STRUT vowels were more subject to accommodation in the task of repeating the male model talker. The FORCE vowel displays the highest degree of convergence in the task of repeating the male model talker, while it is the HAPPY vowel in the case of repeating the female model talker.

Male participants have converged equally towards both model talkers. Vowels displaying a lower difference in duration value when repeating the male model talker are the DRESS, FLEECE, FORCE, GOOSE and NURSE vowels, while those displaying a lower difference in duration value when repeating the female model talker are the HAPPY, KIT, LOT, STRUT and TRAP vowels. Vowels from the GOOSE set display the highest degree of accommodation in the task of repeating the male model talker, while it is those from the HAPPY set in the task of repeating the female model talker.

Previous observations showed that all groups except advanced learners tend to converge more towards the female model talker in terms of vowel duration. The analysis of the difference in duration values across groups for each vowel has demonstrated that advanced learners would converge more towards the male model talker. Intermediate and male learners appear to have converged equally towards both model talkers, and female participants to have converged more towards the female model talker. Overall, the DRESS, FORCE and GOOSE vowels display the highest degree of convergence across groups when repeating the male model talker, and the KIT and LOT vowels the highest degree of convergence across groups when repeating the female model talker. We can conclude that, regarding vowel duration, high vowels appear to be more imitated than low vowels. Another observation that can be made from this analysis is that, overall, vowels displaying negative difference in duration values across groups

are long vowels. This suggests that the mean duration of these vowels when produced by the participants is shorter than when produced by the model talkers. Since there is no dichotomy between short and long vowels in the French phonological system, it appears that participants have difficulty marking the difference existing between short and long vowels in English.

3.5. Creaky voice

While segmental phonetic features seem to display deliberate accommodation, one sociolinguistic feature consistently appears for some French advanced female speakers. Creaky voice – a noted feature in the pronunciation of American speakers (Yuasa, 2008) – is a particular phonation type referring to "a vocal effect produced by a very slow [and irregular] vibration of the vocal cords" (Crystal, 1997: 98), which generates very low pitch (F0), as well as low airflow rates (Podesva, 2013: 429). The idea is to compare the number of creak occurrences between the two tasks to determine whether they are related to linguistic idiosyncrasies or to accommodation towards the native speakers.

Creaky voice is characterised by an irregular and low F0, a discontinuous F0 track and by doubled pulses in wide-band spectrogram. All creaky vowels were then tracked.

Results showed that 3 female advanced learners creaked much more than the others, and more in the reading task, while the native speakers did not, or almost not, creak: one occurrence of creak was observed in the female native speaker's production and four in the male native speaker's production. This was not surprising for creaky voice is a noted feature in the American speakers' pronunciation (Yuasa, 2008), but not so much in the British speakers' pronunciation.

According to the Intrinsic Fundamental Frequency (IF0) theory, low vowels have a lower pitch than high vowels (Whalen and Levitt, 1995). The tongue position required in high vowels pulls on the larynx, which increases the tension on the vocal folds, resulting in a higher F0. Since creaky voice is produced with low longitudinal tension of the vocal folds, it is harder to achieve on high vowels. All vowels on which creak occurred in our study were analysed to observe whether low vowels were more subject to creak. Low vowels appearing at the end of prosodic constituents appeared to be privileged (see Table 17), but one female speaker tended to creak more frequently at the beginning of prosodic constituents.

Table 17. Percentage of creaky voice according to vowel height in both tasks. Remaining percentage corresponds either to creak occurring during the production of a diphthong or a schwa.

Height / Task	Low	High
Reading	53%	21%
Repetition	51%	16%

To conclude, the fact that more occurrences of creak were observable in the reading task indicates both linguistic idiosyncrasies as well as accommodation towards the native speakers. Since the native speakers did not creak and female advanced learners creaked more in the reading than in the repetition task, this means that they may have converged towards the native speakers by creaking less.

4. Discussion and conclusion

While most of the studies regarding the process of phonetic accommodation have been conducted amongst speakers sharing the same language or same dialect, the aim of this research study was to determine whether conclusions drawn mostly from intralanguage studies of spontaneous conversations can also be validated in the case of native/non-native interactions.

We expected female speakers to converge more, whatever their level of proficiency, and level to impact the degree of imitation. Contradictory results were observed: male subjects were more proficient at imitating the native speakers than female subjects. In cases where level was found to impact the degree of accommodation, our hypothesis that advanced learners would converge more than intermediate learners was validated. Usually, perceptual and production skills are strongly correlated with accuracy in producing L2 vowels and with the ratio of L2/L1 usage (Best and Tyler, 2007: 20). We have nevertheless seen that all participants had been grouped into levels based on the number of years spent studying English. Some of the intermediate learners would, however, use English on a regular basis. Hence, the proficiency was not even amongst groups, resulting in a high degree of inter-individual variability.

Our results also corroborate the hypothesis that speakers, whether male or female, converge more towards male than towards female speakers (Bilous and Krauss, 1988, Gallois and Callan, 1988, Willemyns *et al*., 1997). However, this could also be due to the fact that the female model talker produced only five sentences compared to fifteen produced by the male model talker. Fewer vowels were, thus, produced by the female native speaker. Listeners have also been more "in contact" with the voice of the male model talker, which might have led them to more accurately perceive specific characteristics of his voice, resulting in a higher degree of imitation.

Our analysis of vowel duration has demonstrated that non-native speakers produce longer vowels than native speakers, which is correlated to the fact that native speakers speak at a faster rate and, thus, produce shorter vowels. We have observed a general tendency amongst participants to decrease vowel length after auditory exposure to the model talkers, indicating convergence. These observations corroborate previous findings (Zając, 2013; Rojczyk, 2013). Short vowels appeared to be more subject to imitation than long vowels. Since there is no dichotomy between short and long vowels in the French phonological system,

it appears that participants have difficulty marking this difference in English. The durational characteristics of short vowels appear to be, then, easier to imitate.

We expected low vowels to be more imitated than high vowels, especially in the F1 dimension. Depending on the type of interaction, either high or low vowels were more imitated. This hypothesis was, thus, partially validated. However, the analysis of normalized formant values has demonstrated that low vowels displayed more imitation in the F1 dimension, corroborating Babel's (2009) results. A possible explanation as to why low vowels tend to be more subject to imitation than high vowels is that low vowels "are characterized by greater mouth opening and jaw lowering, which leaves more space for individual variability in their production […]. Such variability will contribute to more pronounced convergence effects observed in imitation" (Rojczyk, 2013: 67-68). Another interesting observation is that the position of the FLEECE, KIT and HAPPY vowels were more similar to each other in the repetition task than in the reading task. The FLEECE and HAPPY vowels appeared to be pronounced similarly as the KIT vowel. As stated by Best and Tyler (2007: 18), "a learner's L1 and L2 phonological systems are not completely separate but are instead situated within and encompassing interlanguage". Some L2 sounds can be similar, but not identical, to one L1 sound. A similar L2 phone is realised "in an acoustically different manner than an easily identifiable counterpart in the L1" (Flege, 1987: 59), so discrimination is really poor, and no new phonetic category is formed. Nevertheless, the fact that the KIT vowel was subject to imitation in most of the cases indicates that adults are capable of learning to produce new phones. Flege *et al* (1997) has, indeed, found that experience in English may lead to improvement in the non-native productions of /ɪ/. Furthermore, the GOOSE vowel, and more particularly the NURSE vowel, displayed the highest degree of imitation. These vowels are not present in the French phonological system and, according to the Perceptual Assimilation Model, discrimination is in this case rather good. Asymmetry between two different phonological systems has reflected better discrimination for the exemplars that fail to show a distinct relationship with perceived similarity to the native category (Flege *et al*., 1995). Regarding the production of vowels, it has been hypothesized that adult learners will produce new L2 vowels more accurately than similar L2 vowels because they are more likely to create additional phonetic categories for new (and not similar) vowels (Flege, 1988).

Our last analysis dealt with the study of a specific feature – creaky voice – which is traditionally ignored or rejected outside the perimeter of phonetic features, and is often referred to as a 'paralinguistic' feature. Our observations indicate that some advanced female learners creaked more than the native speakers, and in the reading task, indicating both linguistic idiosyncrasy and accommodation towards the native speakers. Low vowels also seem to be favoured. In addition, creaky voice has been found to be used prosodically to mark the beginning and/or end of a phrase (Gordon and Ladefoged, 2001). Both scenarios were observed in this analysis.

The amount of convergence has been found to be affected by the linguistic distance between participants. Talkers sharing the same language or same dialect tend to converge more than talkers from different linguistic communities (Gambi and Pickering, 2013; Kim *et al*., 2011). Hypotheses tested in this analysis were mostly based on conclusions drawn from spontaneous interactions between people sharing the same language or same dialect. Even if not all of them were validated in the case of deliberate accommodation of L2 speech, the conclusions drawn provide evidence that native and non-native speakers use similar strategies to collapse social distance. In addition, this research study has proved that the different tasks have no effect on imitation in some cases. This can be explained by the high variability phonetic training (Barriuso and Hayes-Harb, 2018). It would have been interesting to design other tasks such as the repetition of an isolated word and/or vowel, and compare the results obtained with those of this analysis.

References

Abercrombie, David. 1967. *Elements of general phonetics*. Chicago: Aldine.

Babel, Molly. 2009. Selective vowel imitation in spontaneous phonetic accommodation. *UC Berkeley Phonology Lab Report*. 163-194.

Babel, Molly. 2011. Imitation in speech. *Acoustics Today* 7(4). 16-22.

Babel, Molly. 2012. Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics* 40(1). 177-189.

Babel, Molly, McGuire, Grant, Walters, Sophia and Alice Nicholls. 2014. Novelty and social preference in phonetic accommodation. *Laboratory Phonology* 5(1). 123-150.

Barriuso, Taylor Anne and Rachel Hayes-Herb. (2018). High variability phonetic training as a bridge from research to practice. *The CATESOL Journal* 30(1). 177-194.

Best, Catherine T. 1995. A direct realist view of cross-language speech perception. In Winifred Strange (ed.), *Speech perception and linguistic experience: issues in cross-language research*, 171-204. Timonium, MD: York Press.

Best, Catherine T. and Michael D. Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In Murray J. Munro and Ocke-Schwen Bohn (eds.), *Second language speech learning: The role of language experience in speech perception and production*, 13-34. Amsterdam: John Benjamin.

Bilous, Frances R. and Robert M. Krauss. 1988. Dominance and accommodation in the conversational behaviours of same- and mixed-gender dyads. *Language & Communication* 8. 183-194.

Black, John.W. (1949). Loudness of speaking: The effect of heard stimuli on spoken responses. *Journal of Experimental Psychology* 39. 311-315.

Boersma, Paul and David Weenink. 2015. Praat: doing phonetics by computer (Version 5.4.08). Available from http://www.praat.org.

Bourhis, Richard Y. 1979. Language in ethnic interaction: A social psychological approach. In Howard Giles and Bernard Saint-Jacques (eds.), *Language and Ethnic Relations*, 117-141. Oxford: Pergamon.

Byrne, Donn. 1997. An overview (and underview) of research and theory within the attraction paradigm. *Journal of Social and Personal Relationships* 14(3). 417-431.

Chartrand, Tanya L. and John A. Bargh. 1999. The Chameleon Effect: The Perception-Behavior Link and Social Interaction. *Journal of Personality and Social Psychology* 76(6). 893-910.

Coupland, Nicolas and Howard Giles. 1988. Introduction: The communicative contexts of accommodation. *Language and Communication* 8(3/4). 175-182.

Eisikovits, Edina. 1987. Sex differences in inter-group and intra-group interaction among adolescents. In Anne Pauwels (ed.), *Women and Language in Australia and New Zealand Society*, 45-58. Sydney: Australian Professional Publications.

Flege, James Emil. 1987. The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics* 15. 47-65.

Flege, James Emil. 1988. The production and perception of foreign language speech sounds. In Harris Winitz (ed.), *Human communication and Its Disorders, A Review. Norwood*, 224-240. NJ: Ablex.

Flege, James Emil. 1995. Second Language Speech Learning: Theory, Findings and Problems. In Winifred Strange (ed.), *Speech perception and Linguistic Experience: Issues in Cross-Language Research*, 233-277. Timomium, MD: York Press.

Flege, James Emil, Munro, Murray J. and Ian R.A. MacKay. 1995. Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America* 97(5). 3125-3134.

Flege, James Emil., Bohn, Ocke-Shwen and Sunyoung Jang. 1997. Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics* 25. 437-470.

Gallois, Cynthia and Victor J. Callan. 1988. Communication accommodation and the prototypical speaker: Predicting evaluations of status and solidarity. *Language and Communication* 8. 271-283.

Gambi, Chiara and Martin J. Pickering. 2013. Prediction and imitation in speech. *Frontiers in Psychology* 4(340). 1-9.

Garnier, Maëva, Lamalle, Laurent and Marc Sato. 2013. Neural correlates of phonetic convergence and speech imitation. *Frontiers in Psychology* 4(600). 1-15.

Giles, Howard. 1973. Accent mobility: A model and some data. *Anthropological Linguistic*s 15. 87-105.

Giles, Howard and Patricia Johnson. 1987. Ethnolinguistic identity theory: A social psychological approach to language maintenance. *International Journal of the Sociology of Language* 68. 69- 99.

Giles, Howard, Coupland, Nikolas and Justine Coupland. 1991. Accommodation Theory: Communication, Context and Consequence. In Justine Coupland, Nikolas Coupland and Howard Giles (eds.), *Contexts of Accommodation: Developments in Applied Sociolinguistics*, 1-68. Cambridge: Cambridge University Press.

Goldinger, Stephen D. 1997. Perception and production in an episodic lexicon. In Keith Johnson and John W. Mullenix (eds.), *Talker variability in speech processing*, 33-66. San Diego: Academic Press.

Goldinger, Stephen D. 1998. Echoes of Echoes? An Episodic Theory of Lexical Access. *Psychological review* 105(2). 251-27.

Goldinger, Stephen D. and Tamiko Azuma. 2004. Episodic memory reflected in printed word naming. *Psychonomic Bulletin & Review* 11(4).716-722.

Kim, Midam, Horton, William S. and Ann R. Bradlow. 2011. Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology* 2(1). 125-156.

Kuhl, Patricia K. and Paul Iverson. 1995. Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *The Journal of the Acoustical Society of America* 97(1). 553-562.

Lakin, Jessica L., Jefferis, Valerie E., Cheng, Clara Michelle and Tanya L. Chartrand. 2003. The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. *Journal of Nonverbal Behavior* 27. 145-161

Laver, John and Peter Trudgill. 1979. Phonetic and linguistic markers in speech. In Klaus Scherer and Howard Giles (eds.), *Social markers in speech*, 1-10. Cambridge: Cambridge University Press.

Laver, John. 1989. Cognitive science and speech: A framework for research. In Helmut Schnelle and Niels O. Bernsen (eds.), *Logic and linguistics: Research directions in cognitive science: European perspective*, vol. 2, 37-71. Hillsdale, NJ: Lawrence Erlbaum.

Namy Laura L., Nygaard, Lynne C. and Denise Sauerteig. 2002. Gender differences in vocal accommodation: The role of perception. *Journal of Language and Social Psychology* 21(4). 422-43.

Natale, Michael. 1975. Convergence of Mean Vocal Intensity in Dyadic Communication as a Function of Social Desirability. *Journal of Personality and Social Psychology* 32(5). 790-804.

Nilsenová, Marie and Marc Swerts. 2012. Prosodic Adaptation in Language Learning. In Jesús Romero-Trillo (ed.), *Pragmatics and Prosody in English Language Teaching*, 77-94. Dordrecht: Springer Science+Business Media B.V.

Pardo, Jennifer S. 2006. On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America* 119(4). 2382-2393.

Pardo, Jennifer S., Cajory Jay, Isabel and Robert M. Krauss. 2010. Conversational role influences speech imitation. *Attention, Perception, & Psychophysics* 72(8). 2254-2264.

Pardo, Jennifer S., Cajori Jay, Isabel, Hoshino, Risa, Hasbun, Sara Maria, Sowemimo-Coker, Chantal and Robert M. Krauss. 2013. Influence of Role-Switching on Phonetic Convergence in Conversation. *Discourse Processes* 50(4). 276-300.

Rojczyk, Arkadiusz. 2013. Phonetic imitation of L2 vowels in a rapid shadowing task. In John Levis and Kimberly LeVelle (eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference*, 66-76. Ames, IA: Iowa State University.

Sato, Marc, Grabski, Krystyna, Garnier, Maëva, Granjon, Lionel, Schwartz, Jean-Luc and Noël Nguyen, N. 2013. Converging towards a common speech code: imitative and perceptuo-motor recalibration processes in speech production. *Frontiers in Cognitive Sciences* 4(422). 1-14.

Tortel, Anne. 2008. ≪ ANGLISH ≫. *TIPA, Travaux interdisciplinaires sur la parole et le langage* 27. 111-122.

Tortel, Anne. 2013. Prosody in a contrastive learner corpus. In Ana Diaz-Negrillo, Nicolas Ballier and Paul Thompson (eds.), *Automatic Treatment and Analysis of Learner Corpus Data*, 227-249. Amsterdam : John Benjamins. B.V

Watzlawick, Paul, Beavin Bavelas, Janet and Don D. Jackson. 1967. *Pragmatics of Human Communication*. New-York: Norton.

Willemyns, Michael, Gallois, Cynthia, Callan, Victor J. and Jeffery Pittam. 1997. Accent accommodation in the job interview: Impact of interviewer accent and gender. *Journal of Language and Social Psychology* 16. 3-22.

Wells. John C. 1982. *Accents of English: The British Isles*, vol. 2. Cambridge: Cambridge University Press.

Welkowitz, Joan, Feldstein, Stanley, Finkelstein, Markand Lawrence Aylesworth. 1972. Changes in vocal intensity as a function of interspeaker influence. *Perceptual and Motor Skills* 35. 715-18.

Zuengler, Jane. 1982. Applying accommodation theory to variable performance data in L2. *Studies in Second Language Acquisition* 4. 181-192.

Zając, Magdalena. 2013. Phonetic imitation of vowel duration in L2 speech. *Research in Language* 11(1). 19-29.

Appendix

Table 2. Statistical analyses resulting from the graphic interpretation of vowel duration per level in the reading task. Mean duration is displayed in millisecond.

Vowel type	Mean duration (advanced)	Mean duration (intermediate)	*p*-value (*t*-test)
NURSE	114.918	113.167	0.41
START	131.1667	122.393	0.06
FLEECE	95.22	90.189	0.16
FORCE	110.452	107.487	0.71
HAPPY	79.343	93.605	< 0.001
TRAP	80.754	86.212	< 0.05
KIT	57.789	68.576	< 0.001
FOOT	59.25	65.641	< 0.05
STRUT	66.278	79.152	< 0.001
GOOSE	94.706	96.441	0.38

Table 3. Statistical analyses resulting from the graphic interpretation of vowel duration per level in the repetition task. Mean duration is displayed in millisecond.

Vowel type	Mean duration (advanced)	Mean duration (intermediate)	*p*-value (*t*-test)
NURSE	108.5	102.5	0.84
START	171.086	170.083	0.55
TRAP	73.266	80.657	< 0.01
DRESS	80.25	90.646	< 0.001
HAPPY	70.845	88.125	< 0.001
KIT	54.947	65.117	< 0.001
LOT	86.455	92.943	< 0.05
FOOT	53.25	73.25	< 0.001
STRUT	56.00	65.889	< 0.001
FLEECE	94.25	100.252	0.07
FORCE	110.65	112.677	0.34
GOOSE	119.417	113.167	0.76

Table 4. Statistical analyses resulting from the graphic interpretation of advanced learners' production between the two tasks. Mean duration is displayed in millisecond.

Vowel type	Mean duration (read)	Mean duration (repetition)	*p*-value (*t*-test)
START	131.167	171.083	< 0.001
GOOSE	94.706	119.417	< 0.05
TRAP	80.754	73.266	< 0.01
HAPPY	79.344	70.85	< 0.01
KIT	57.789	59.947	< 0.01

Vowel type	Mean duration (read)	Mean duration (repetition)	*p*-value (*t*-test)
FOOT	59.25	53.25	< 0.05
STRUT	66.278	56.00	< 0.001
NURSE	114.918	108.50	0.19
FLEECE	95.22	94.25	0.42
FORCE	110.452	110.65	0.51
LOT	86.455	89.867	0.14

Table 5. Statistical analyses resulting from the graphic interpretation of intermediate learners' production between the two tasks. Mean duration is displayed in millisecond

Vowel type	Mean duration (read)	Mean duration (repetition)	*p*-value (*t*-test)
START	122.393	170.083	< 0.001
FLEECE	90.19	100.251	< 0.01
GOOSE	96.441	113.167	< 0.01
FORCE	107.487	112.677	0.14
FOOT	65.641	73.25	0.08
NURSE	113.167	102.50	< 0.05
STRUT	79.152	65.889	< 0.001

Table 6. Statistical analyses resulting from the graphic interpretation of vowel duration per gender in the reading task. Mean duration is displayed in millisecond.

Vowel type	Mean duration (female)	Mean duration (male)	*p*-value (*t*-test)
TRAP	87.035	79.899	< 0.05
NURSE	121.333	106.885	< 0.05
START	134.957	118.917	< 0.01
DRESS	98.716	83.354	< 0.001
HAPPY	90.533	82.406	< 0.05
KIT	66.043	60.212	< 0.001
FLEECE	101.887	83.481	< 0.001
GOOSE	102.712	88.487	< 0.01
STRUT	77.231	68.167	< 0.01
LOT	97.315	86.346	< 0.001
FOOT	69.231	55.75	< 0.01
FORCE	112.864	105.07	0.93

Table 7. Statistical analyses resulting from the graphic interpretation of vowel duration per gender in the repetition task. Mean duration is displayed in millisecond.

Vowel type	Mean duration (female)	Mean duration (male)	*p*-value (*t*-test)
TRAP	81.10	72.741	< 0.01
NURSE	111.5	99.5	< 0.05
DRESS	91.895	79.00	< 0.001
HAPPY	81.937	77.075	< 0.05
KIT	62.093	57.885	< 0.001
FLEECE	104.625	89.811	< 0.001

Vowel type	Mean duration (female)	Mean duration (male)	*p*-value (*t*-test)
STRUT	64.611	57.278	< 0.001
LOT	93.278	86.12	< 0.05
FOOT	68.25	58.25	< 0.05
GOOSE	122.083	110.50	0.1
FORCE	112.864	105.075	0.07

Table 8. Statistical analyses resulting from the graphic interpretation of female learners' production between the two tasks. Mean duration is displayed in millisecond.

Vowel type	Mean duration (read)	Mean duration (repetition)	*p*-value (*t*-test)
TRAP	87.035	81.1	< 0.05
DRESS	98.716	91.896	< 0.01
HAPPY	90.533	81.936	< 0.01
KIT	66.043	62.093	< 0.001
STRUT	77.232	64.611	< 0.001
NURSE	121.33	111.50	0.07
LOT	97.315	93.278	0.1
FOOT	69.231	68.25	0.4
FLEECE	89.811	104.625	< 0.001
START	169.167	172.00	0.36
FORCE	105.075	112.864	0.07
GOOSE	110.50	122.083	0.09

Table 9. Statistical analyses resulting from the graphic interpretation of male learners' production between the two tasks. Mean duration is displayed in millisecond.

Vowel type	Mean duration (read)	Mean duration (repetition)	*p*-value (*t*-test)
TRAP	78.899	72.741	< 0.01
DRESS	83.354	79.00	< 0.05
HAPPY	82.406	77.075	< 0.05
KIT	60.212	57.885	< 0.05
STRUT	68.167	57.278	< 0.001
NURSE	106.885	99.50	0.12
LOT	86.345	86.12	0.47
START	118.917	169.167	< 0.001
GOOSE	88.487	110.50	< 0.001
FLEECE	83.481	89.811	0.06
FORCE	105.075	110.00	0.17
FOOT	55.75	58.25	0.27

Table 10. Statistical analyses resulting from the graphic interpretation of advanced learners' production between the two tasks of repetition. FNS corresponds to the task of repeating the female native speaker and MNS to the task of repeating the male native speaker. Mean duration is displayed in millisecond.

Vowel type	Mean duration (FNS)	Mean duration (MNS)	*p*-value (*t*-test)
TRAP	63.625	79.748	< 0.001
DRESS	74.5	81.4	< 0.05
LOT	74.454	94.389	< 0.001
GOOSE	87.25	135.50	< 0.001
FLEECE	84.167	100.30	< 0.01
FORCE	117.0	104.3	< 0.05
HAPPY	85.80	63.991	< 0.001
NURSE	111.0	103.5	0.2
KIT	55.961	54.334	0.1
STRUT	56.125	55.90	0.47

Table 11. Statistical analyses resulting from the graphic interpretation of intermediate learners' production between the two tasks of repetition. FNS corresponds to the task of repeating the female native speaker and MNS to the task of repeating the male native speaker. Mean duration is displayed in millisecond.

Vowel type	Mean duration (FNS)	Mean duration (MNS)	*p*-value (*t*-test)
TRAP	70.38	87.479	< 0.001
DRESS	78.125	93.150	< 0.001
LOT	80.672	101.056	< 0.001
GOOSE	92.50	123.50	< 0.001
FLEECE	85.085	109.20	< 0.001
NURSE	94.50	118.50	< 0.01
STRUT	63.50	67.80	0.1
FORCE	129.40	95.612	< 0.001
HAPPY	102.00	81.818	< 0.001

Table 12. Statistical analyses resulting from the graphic interpretation of female learners' production between the two tasks of repetition. FNS corresponds to the task of repeating the female native speaker and MNS to the task of repeating the male native speaker. Mean duration is displayed in millisecond.

Vowel type	Mean duration (FNS)	Mean duration (MNS)	*p*-value (*t*-test)
TRAP	71.125	87.750	< 0.001
DRESS	81.250	94.025	< 0.001
LOT	78.908	102.778	< 0.001
GOOSE	90.75	137.75	< 0.001
FLEECE	91.333	112.60	< 0.001
NURSE	107.25	120.00	0.07
FORCE	124.00	102.525	< 0.001
HAPPY	95.60	75.727	< 0.001
STRUT	64.625	64.600	0.5

Table 13. Statistical analyses resulting from the graphic interpretation of male learners' production between the two tasks of repetition. FNS corresponds to the task of repeating the female native speaker and MNS to the task of repeating the male native speaker. Mean duration is displayed in millisecond.

Vowel type	**Mean duration (FNS)**	**Mean duration (MNS)**	***p*-value (*t*-test)**
TRAP	62.784	79.407	< 0.001
DRESS	71.375	80.525	< 0.01
LOT	76.218	92.667	< 0.001
GOOSE	89.00	121.25	< 0.001
FLEECE	77.796	96.90	< 0.001
NURSE	98.25	102.00	0.32
STRUT	55.00	59.10	0.08
FORCE	122.40	97.475	< 0.001
HAPPY	92.20	70.137	< 0.001

NON-NATIVE ATTITUDES TO /θ/ AND /ð/: A EUROPEAN CASE STUDY*

RIAS VAN DEN DOEL
Utrecht University, the Netherlands
w.z.vandendoel@uu.nl

ANNE-FRANCE C.H. PINGET
Utrecht University, the Netherlands
a.c.h.pinget@uu.nl

HUGO QUENÉ
Utrecht University, the Netherlands
h.quene@uu.nl

Abstract
This paper investigates the evaluation of the English sounds /θ/ and /ð/ as produced by European non-native speakers. Using the data from a larger web survey, we compared the error judgements by different native and non-native users of English. This was done to establish whether there is any normative convergence among European non-native speakers, or if this was counteracted by other patterns, such as the presence or absence of these sounds in their L1s. Our analysis shows that while European non-native judges do not differ consistently from native-speakers in their judgements, there are also subtle differences between different groups of non-native speakers, implying that we should be careful not to generalise across groups about non-native attitudes to these sounds.

Keywords: attitudes, dental fricatives, endonormative, Euro-English, exonormative, non-native speakers

1. Introduction

Due to increasing globalisation, more communities are exposed to linguistic variation, putting them in a position to evaluate the many different language varieties and features they encounter. According to Blommaert (2001:126), this "has opened a new and wider space for measuring diversity *as aberrance* from newly reinforced or reinvented standards, customs and benchmarks. . . . Rejection and exclusion on the grounds of features of one's linguistic resources are rife, although such reactions need not always be categorical." The example Blommaert (2001: 81) provides of "the production of local, deviant normativity" by non-

* We gratefully acknowledge the use of the Speech Accent Archive under the Creative Commons License.

native speakers (NNSs) of English in a South African township can illustrate how different, localised normative practices are emerging globally. Interestingly, Blommaert's description of a locally differentiated normativity is conceptually quite distinct from the idea of ever-converging norms induced by the shared interests and practices of larger communities (such as NNSs of English) as proposed by some other scholars. The notion of normative convergence is, for instance, implicit in early conceptualisations of a pan-European English by English as Lingua Franca (ELF) researchers such as Seidlhofer (2001). The latter has argued that Europe is "witnessing the emergence of an endonormative model of lingua franca English which will increasingly derive its norms of correctness and appropriacy from its own usage rather than that of the UK or the US" (2001:15). Another ELF researcher, Jenkins, has also addressed the issue of normative convergence within this context:

> For example, only two mainland European languages, Greek and Spanish, include the physiologically difficult sounds /θ/ and /ð/ in their pronunciation repertoires. . . . When they speak English, the majority of Europeans substitute these sounds with either /t/ and /d/ or /s/ and /z/. It is therefore unlikely that /θ/and /ð/ will be features of "Euro-English" accents. What is not clear at this early stage is whether the former – as used by many Italian and Scandinavian speakers of English, or the latter – as used by many French and German speakers of English, will ultimately become *the accepted norm*, or whether there will be scope for regional variation in this respect within "Euro-English" (2001: 17, our italics).

Even though Jenkins (2017: 343) has distanced herself from some of the views expressed in this article, it may still be interesting to explore European NNSs' normative attitudes. The rationale for this would not be merely to establish whether ELF researchers are justified in disowning any previous claims about European English. If any patterns are found in the assessment of specific features, this would help to support or reject categorical claims about the evaluative behaviour of users of English in terms of convergence or differentiation. We may, for instance, find that specific groups of NNSs do indeed agree on the relative insignificance of particular features of non-native speech. Alternatively, we may be able to uncover distinctly local patterns in NNSs' "production" of normativity – ranging from a preference for a nativised pronunciation model to a closer alignment with NS norms (idealised or otherwise), or even to a more truncated repertoire in accent evaluation (cf. Blommaert 2011: 213). If such differentiation is indeed attested, it would reflect the complex responses found in various groups to language features to which they are exposed through globalisation, in addition to any local variation in exposure to English and in levels of proficiency.[1]

The English fricatives /θ/ and /ð/, labelled as dental in a British and as interdental in a North American context (Ladefoged and Maddieson 1997: 143–

[1] Since the differential effects of proficiency on normative convergence do not feature largely in ELF conceptualisations of a pan-European English, we have not addressed the issue here.

144; but see Collins and Mees 1996: 141–2), are often described as difficult to master for NNSs. Despite their occurrence in a wide range of different languages, including varieties of Spanish, Greek, Arabic, Burmese and Swahili, these sounds are widely viewed as exotic or "marked" (Wells 2010), and this is seen as a factor in inhibiting their acquisition (Derwing and Munro 2015: 70). Seen as indexical of NS accents, many NNSs do indeed consider /θ/ and /ð/ (hereafter also referred to as dental fricatives, or DFs) to be salient potential stumbling-blocks – an added difficulty for some being the confusion arising from the representation of both sounds as the same digraph <th> (Collins and Mees 1996: 142).

DFs are often reported to be absent from many inner and outer circle varieties of English. Often, however, the substitution of either /θ/ or /ð/ by other sounds is dependent on phonological, semantic, pragmatic and sociophonetic contexts (cf. Schneider 2004: 1123). For instance, while /θ/ may be realised in Southern Irish English as a dental or an alveolar stop, some speakers may use [θ] word-finally "in careful speech or in reading style" (Hickey 2004: 330). Similarly, even though neither sound is used in Jamaican Creole, they both feature in more acrolectal Jamaican English (Devonish and Harry 2004: 477). This suggests that categorical claims about the absence of DFs from specific varieties (such as those in Walker 2010: 29) should not be taken at face value. This is not a trivial point, since the assumption that /θ/ or /ð/ do not occur in many varieties of English has contributed to the perception of their reduced relevance to learners (cf. Crystal 2001: 57). For instance, Kirkpatrick states:

> I feel sorry for poor learners of English who spend hours of classroom time trying to master the R[eceived] P[ronunciation] sounds of /θ/ and /ð/, as these are difficult sounds to learn if they do not exist in your language and, it turns out, they are not used in many varieties of English anyway (2007: 17).

Jenkins (2000: 138) has also argued that it is unnecessary for NNSs to acquire sounds such as the DFs that do not exist in all NS varieties. In her view, it is "unreasonable to have 'higher' expectations of non-native as opposed to native speakers" (p. 139). In addition, Jenkins (2000: 137) has claimed to have empirical evidence that DFs are irrelevant to intelligibility in lingua franca English. This is why Jenkins did not include these sounds in her well-known list of essential non-native pronunciation targets, the Lingua Franca Core or LFC (2000: 159ff.). While the LFC does not actually proscribe /θ/ and /ð/ in non-native English, their use may well be inappropriate in specific contexts. As Jenkins has pointed out, some native speakers (NSs) "will have to accept that it may already be better, depending on their E[nglish] as an I[nternational] L[anguage] interlocutor, to use substitutions of /θ/ and /ð/" (2000: 228). Anyone insisting that the LFC is exclusively concerned with intelligibility may be struck by the evaluative overtones of such pronouncements, which appear to be concerned with norm-based appropriacy.

Even though ELF researchers have moved away from their earlier interest in features (Jenkins et al. 2011: 288), this does not mean that all the suggestions of the LFC have been abandoned, such as the recommendation to drop DFs from the pronunciation curriculum (e.g. Walker 2010). This suggestion has even been adopted by vociferous critics of the LFC, such as MacKenzie (2014: 126), who also stresses the absence of DFs in many NS and NNS varieties of English. However, MacKenzie also recognises that didactic considerations should not only be based on "mere intelligibility", but also take account of NNSs' own attitudes and aspirations (2014: 132–3) – a point previously made by Van den Doel (2008). This is important, because the LFC's overriding concern with intelligibility may obscure any issues non-native users have with the acceptability of particular non-native pronunciations. If, as Jenkins has posited (2000: 137), non-native realisations of the DFs do not affect intelligibility in lingua franca English, any persistence on the part of NNSs in evaluating these as "errors" suggests that considerations of acceptability also play a part in their normative behaviour. It may even imply that such NNSs have become more responsive to the stigmatisation which which different groups of NSs tend to regard NNS realisations of /θ/ and /ð/ (Van den Doel 2006: 290; Jenkins 2000: 138), and in some cases actively contribute to such stigmatisation.

While ELF researchers tend to be more concerned with intelligibility than acceptability, and may to some extent have abandoned their earlier interest in endonormative convergence in European English, the issue of local European norms and varieties has continued to interest a few scholars. Apart from general studies on Euro-English (Mollin 2006), attention has been paid to lexico-grammar (Breiteneder 2009; Forche 2012), pragmatics (Klimczak-Pawlak 2014), and to describing specific European varieties of English (Bushfeld 2011; Salakhyan 2012; Kautzsch 2014; Edwards 2016). Apart from Jenkins's interesting suggestions about the pronunciation features of European English (2001:17), phonology has rarely been discussed. Without explicitly referring to the notion of European English, Beinhoff (2008, 2013) has investigated Greek and German listeners' attitudes to Greek and German accents in English. In addition, Van den Doel and Quené (2013) have investigated claims of emerging phonological norms in European non-native speakers (Eu-NNSs), but without considering individual sounds.

Recently, Modiano has drawn renewed attention to the notion of Euro-English, speculating that Brexit "will clear the sociolinguistic space for the emergence of an authentic European English" (2017: 313). The mixed reactions to his article suggest that the issue continues to be controversial, and could benefit from fresh scholarly attention. Of course, we may well want to be content with Schneider's claim that "empirical, realistic linguists . . . have consistently failed to identify such a variety", and that there is "no evidence for a homogenizing tendency likely to produce a single, reasonably coherent variety in the long run" (2017: 353). Admittedly, it may be difficult to produce evidence for the notion of European English as a "reasonably coherent variety." But an examination of Europeans'

attitudes to non-native English may still reveal homogenizing tendencies pointing to some kind of normative convergence. They may also attest to the distinctly local patterns of differentiation proposed by Blommaert (2011).

For instance, it would be interesting to know how European non-native speakers of English (Eu-NNSs) identified, prioritised and evaluated non-native realisations of two sounds which, although often considered to be indexical of NS accents, have been reported to be irrelevant to intelligibility. If Eu-NNSs as a group are indeed more accepting of any substitutions of DFs by other Europeans than their NS (or possibly non-European non-native) counterparts, such a more accommodating orientation to the NNS accents of their fellow Continentals may be interpreted in different ways. It could be variously seen as evidence of endonormative convergence, however limited, among Eu-NNS, or perhaps even of a pragmatic indifference, shared with other NNSs, to the preservation of non-essential phoneme contrasts. However, if such relative leniency is not attested, this would suggest that Eu-NNSs' judgements of NNS realisations of /θ/ and /ð/ are affected by factors other than "mere intelligibility" – such as acceptability. If, for instance, some NNSs have become susceptible to NS stigmatisations of NNS realisations of DFs, this would be of interest to those wishing to review any ELT pronunciation training curricula designed to serve such users' interests.

Since some Eu-NNSs actually have /θ/ and /ð/ in their phonological inventories (e.g. speakers of Greek and Castilian Spanish), it would be interesting to explore if speakers of these languages identify, prioritise and evaluate DF-substitutions any differently from judges whose languages do not feature any DFs at all (e.g. Dutch, Polish and Finnish).[2] This would reveal whether, in the case of these sounds, local norms override any pan-European normative convergence. It has already been shown that NS judgements of NNS realisations of /θ/ and /ð/ are affected by the way these sounds were realised in the accents the judges were familiar with (Van den Doel 2006). If such "accent parallelism" between linguistic background and non-acrolectal realisations also impacts NNSs' judgements of the speech of other NNSs, this would suggest that transferred L1 norms, rather than any considerations of the non-nativeness of the speech judged, play a part in lingua franca communication. This has not been investigated systematically with regard to DFs, but a precursor study by Beinhoff (2008) revealed that German and Greek listeners evaluated non-native realisations of /θ/ equally severely.

Given the discussion about NNSs' responses to NNS realisations of /θ/ and /ð/ in especially a European context (e.g. Jenkins 2001), we have attempted to determine if, and possibly to what extent, Eu-NNS listeners demonstrate any evaluative convergence in their identification, prioritisation and evaluation of these sounds as pronounced by fellow NNSs. In order to investigate this, we analysed additional data from the Internet survey described by Van den Doel and

[2] For reasons of space, we have omitted any references to the allophonic and sociolinguistic variation of DFs in Spanish. For a concise overview of allophonic variation, see MacKenzie n.d.).

Quené (2013), which was designed to enable the assessment of speech samples produced by different NNSs of English from the European continent.

2. Methodology

The data used for our analysis are derived from an existing Internet survey (Van den Doel and Quené, n.d.), the details of which will only be discussed here where relevant to the present study. For an exhaustive description, see Van den Doel and Quené (2013). The survey was set up so as to allow European NNS speech samples to be judged by three different groups of respondents: (1) Eu-NNSs, drawn from Croatia, Greece, Poland, Spain and the Netherlands; (2) NSs of English from different Inner Circle countries; (3) non-European non-native speakers (NEu-NNSs) from the People's Republic of China. This made it possible to compare the judgements of the different groups, and interpret clear internal consistency within Group (1), or any significant divergences in evaluation between Group (1) and either Groups (2) or (3), as indications of any emerging Continental European norms for accent evaluation. Unlike in the precursor study, we will only be concerned with the three groups' identification, prioritisation and assessment of non-native realisations of DFs – where necessary in relation to the evaluation of other features.

As in Van den Doel (2006), the Internet survey employed was designed to be an accessible platform for both solicited and unsolicited respondents in different locations to download and listen to sound files of different NNSs, and to facilitate the selection and evaluation of any pronunciation features in the stimuli which respondents choose to identify as errors. The survey was kept deliberately short and simple in order to attract respondents with diverging interests and educational backgrounds, featuring simple instructions on pronunciation assessment and no more than 30 audio stimuli. These consisted of three different sentences read by two speakers (male and female) of five Continental European languages, only two of which, Greek and Castilian Spanish, included DFs in their phoneme inventories. The sentences used as audio stimuli had been taken from a larger reading passage entitled "Please Call Stella", as recorded by both NSs and NNSs at the Speech Accent Archive (Weinberger 2011), and had been selected on the basis of their potential for generating salient non-native accent features – DF substitutions in particular. The sentences included in the survey, which contained six tokens eligible for DF-substitution (in bold), were:

(1) And maybe a snack for her bro**th**er Bob.

(2) Ask her to bring **th**ese **th**ings wi**th** her from **th**e store.

(3) Five **th**ick slabs of blue cheese.

For most speakers of British Received Pronunciation, these sentences would feature two instances of initial /θ/ (*things, thick*), two of initial /ð/ (*these, the*), one

of medial /ð/ (*brother*), and one of final /ð/ *(with*). Note that a minority of speakers of educated British English and some other varieties would use /θ/ in *with*, as would a majority of speakers of General American and Scottish English (Wells 2008: 904). Of course, it would have been preferable to use a word that consistently has final /ð/ in most reference varieties of English (such as *breathe*), but this was unfortunately not included in the relevant reading passage. Similarly, tokens involving medial and final /θ/ were also missing.

The non-native English accents included in the stimuli had been selected to represent major European language groups (Castilian Spanish for Romance, Dutch for Germanic, Finnish for Finno-Ugric, Greek for Hellenic, and Polish for Slavic) and to showcase diversity with the regard to the inclusion of /θ/ and /ð/. Auditory analysis by an expert NS phonetician showed that all Greek and Spanish speakers included in the study consistently pronounced all DFs, whereas this was not the case for the Dutch, Polish and Finnish speakers (whose L1s do not feature these sounds). It would have been ideal if the Dutch, Polish and Finnish speakers had also demonstrated more varied patterns of DF-substitution, but virtually no patterns other than **th**-stopping were attested. Of course, additional speaker variation in DF-substitution could have made the analysis possibly less reliable. Even so, there was already considerable variation between speakers, since perfectly matched guises are not easily produced, or indeed available from the Speech Accent Archive. Partly because respondents objected to the duration of an earlier pilot, we decided only to use a limited number of verbal guises with roughly comparable levels of proficiency. We considered this to be appropriate, since the focus of the experiment is on listeners' possibly converging attitudes to specific features found in the same speakers, rather than on the speakers' performance itself.

We also decided, in line with previous experiments of this kind, to ask respondents to identify any non-native realisations as "errors". While this may predispose respondents to judge the stimuli from an overly prescriptive or NS perspective, the use of an error-based framework is likely to make the experiment more accessible to non-linguists. If the experiment had been preceded by a discussion about the relative arbitrariness of errors within the context of English as a European lingua franca, this could have confused and biased any potential respondents as well. Consequently, we introduced the concept of European Englishes on the experiment's welcome page in fairly neutral terms, with a focus on intelligibility. In addition, since it took extra effort to identify any errors, the set-up of the experiment implicitly encouraged respondents not to designate any deviations from NS norms as erroneous. Thus, it could be argued that the number of errors identified by respondents is a direct reflection of their commitment to participating in the experiment. It could even be an indication that they were especially concerned with factors such as acceptability and stigmatisation.

As is described in Van den Doel and Quené (2013), the Speech Accent Archive recordings were edited, downsampled and each presented, in a random order, on separate web pages of a specially designed web survey. Apart from being asked

to provide a global assessment of the way each individual sentence had been pronounced, respondents were given the opportunity to identify zero to three errors in each of the utterances, by clicking on orthographic representations of *all* the phonemes which could conceivably be assessed as incorrectly pronounced, and to indicate the gravity of such errors, before being allowed to go on to the next web page (see Figure 1). Thus, for each stimulus, it was possible to record (1) participants' global evaluation of the utterance (not considered in the present article); (2) the number and nature of the errors they believed they heard; (3) the severity assigned to each of these errors. Global evaluations (1) were obligatory, but error identification (2) and severity rating (3) were not.

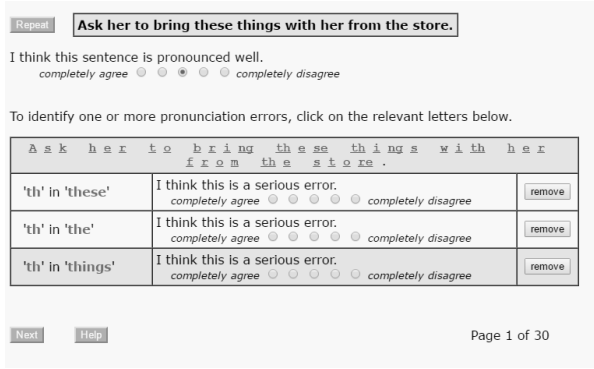

Figure 1. A sample page from the Internet survey held by Van den Doel and Quené (n.d.)

The survey, available without a password at let.uu.nl/~Rias.vandenDoel/personal/wwstim/eureng/html/, was not specifically targeted at any groups, but since respondents had been approached primarily through the authors' academic network, an educational bias may have been likely (see Van den Doel and Quené 2013: 83). Some respondents, however, had been approached through social media, and all had been offered the opportunity to take part in a lottery for a small prize as an incentive. A more detailed description of the experiment may be found in Van den Doel and Quené (2013: 80–83).

Between 2 April 2009 and 25 April 2010, the responses of 373 participants were collected (see Van den Doel and Quené 2013). The analysis provided in this paper is only based on the responses of a subset of the whole respondent pool (n=350), because 23 participants did not identify any phonemes in the error identification task. Respondents in this subset consisted of (1) 279 self-identified NSs of Croatian (n=22), Dutch (n=121), Greek (n=28), Polish (n=89) and Spanish (n=19); (2) 40 self-identified NSs of English; and (3) 31 self-identified NSs of Chinese, i.e. NEu-NNSs of English. No additional information was logged about the language backgrounds of the NSs of English and of Chinese, making it impossible to compare and contrast the judgements of British and other NSs of

English (as in Van den Doel 2006), or to investigate any differences between Chinese languages or dialects.

All selected responses were subjected to multi-level modelling (Kreft and De Leeuw 1998; Luke 2004; Quené and Van den Bergh 2008). Each model took into account the variances between judges, between speakers, and between items. This implies that the resulting regression coefficients are "corrected" for random variation between and within judges. The multi-level analyses in this study were all performed with the R programming environment. Computations and evaluations were carried out with functions from the packages *lme4* (Bates et al. 2015) for R (R Development Core Team 2016).

In this way, two dependent variables were modelled:

i. The *hit rate*, i.e. the probability of a sound being reported as an error by the judges in question. This was not only estimated for all sounds (*overall hit rate*), but also for all DFs only (*DF hit rate*). If any group's DF hit rate approximated the NS baseline value (intercept), this would mean that the number of DF errors identified by the listeners was similar to that selected by the NS. Additionally, if the proportion of DFs actually designated as errors by any group is compared to the proportion of all sounds being reported as erroneous, the resulting *relative hit rate* should indicate the prioritisation given to the selection of DFs as opposed to other potential errors.

ii. The *error severity estimate*, or the degree (on a 5-point Likert scale) to which judges agreed with the statement "I think this is a serious error" with reference to any sound they had reported as erroneous. This can be calculated for all sounds (*overall error severity estimate*), but also for the DFs (*DF error severity estimate*). If the latter estimate is significantly higher or lower than the NS baseline value (intercept), this would signal a consistent deviation from NS norms in terms of the perceived gravity of the errors selected. As with the hit rates, it is also possible to compare the overall and the DF error severity estimates, in order to determine if any group evaluated DF errors significantly more or less strictly than other identified errors. This *relative error severity estimate* would be another important indicator of the priority given to DFs as opposed to other sounds by specific groups. This estimate, however, refers to the strictness with which the DFs were assessed, rather than to their selection as errors.

We felt the need to distinguish clearly between error identification and assessment, simply because judges sometimes report a great many errors which they consider unimportant, or report fewer errors than do other groups but assess these more critically. For instance, Van den Doel's (2006) investigation of NS evaluation of Dutch-accented pronunciation features revealed how British judges in particular tended to over-report certain errors while simultaneously denying their significance, and how North American respondents would proffer stricter evaluations of the lower number of features they judged to be erroneous. Van den Doel (2006) speculates that such trends may be informed by underlying attitudes

to errors as either "noticeable but not serious" or "serious only where noticeable" (see Van den Doel 2006: 297 for details). Interestingly, the North American tendency to be stricter about fewer errors was also attested for a number of DF-substitutions (p. 297) – a useful reminder that there may be considerable disagreement between different groups of NSs evaluating non-native realisations of DFs.

Because of differences between groups of NSs in the evaluation of DFs, we should not ascribe too much authority to individual NSs' judgements of non-native realisations of these. An estimate of the judgements given by a larger, varied group of NS respondents may be a more reliable indicator of the degree of approximation to acrolectal NS realisations of the DFs, and thus be used as a basis to compare and contrast the potentially different levels of divergence from these responses by both Eu-NNSs and NEu-NNSs. It is for this reason, and not with a view to prioritising NS practices or norms, that we decided to employ a NS baseline or intercept. This is also warranted by our research objectives, which are concerned with establishing significant patterns of variation in error evaluation among groups of judges, rather than with the actual performance of the speakers and the inevitable differences between them.

3. Results

3.1. Identification

In order to calculate the probability of a DF realisation being reported as an error by the different groups of judges, we investigated the hit rate. If a particular judge had marked such a realisation as an error, then we noted this as a "hit". These binary data (hit/miss) were analysed by means of mixed-effects logistic regression (GLMM, Quené and Van den Bergh 2008). The dependent variable in this analysis is based on the odds ratio of observing a hit: if the proportion of hits is $P=0.8$, then the corresponding odds ratio is $(P/(1-P))=4$. For computational purposes, we worked with the natural logarithm of this odds ratio, $\ln(P/(1-P))=1.39$, rather than with the odds ratio itself. The logistic regression model attempts to estimate the "log-odds" or "logit" of a hit (Hosmer and Lemeshow 2000; Pampel 2000), taking language background as a fixed predictor, and the variance between judges, speakers and between items as random factors. As was noted in Section 2, the NSs of English were used as the baseline group. In other words, the regression coefficient reported for the NS group constitutes the intercept or baseline, and the coefficients reported for the other groups constitute deviations (positive or negative) relative to this baseline. The mixed-effects logistic regression analysis of the DF hit rate has been summarised in Table 1.

Table 1. Estimated coefficients for log odds of the DF hit rate (marking a DF as an error), broken down by respondents' language backgrounds. For fixed effects, the regression coefficients are given; for random effects, their standard deviations are given

Fixed effects	Coefficient	(SE)	Z	P
NS (baseline)	−2.598	(0.389)	−6.684	<.001
Eu-NNSs	−0.467	(0.154)	−0.304	.761
NEu-NNSs	−0.413	(0.223)	−1.853	.064
Random effects	Std. Dev.	N		
Judges	0.763	350		
Speakers	0.770	10		
Sentences	0.460	3		
Residuals		21004		

The coefficients reported show that the estimated log odds of a hit for the baseline group of judges, the self-declared NSs of English, was −2.598, corresponding to an observed hit rate of 0.093 (thus 9.3% of the DF realisation marked as an error). The probability of marking a DF realisation as an error for the Eu-NNS judges (estimated log odds −2.598−.467, observed hit rate 0.093) was not significantly different from the NSs. The NEu-NNS (Chinese) judges marked fewer DF realisations as errors than did the NS group (estimated log odds −2.598−0.413, observed hit rate 0.067), which was only marginally significantly different from the NSs. This implies that there are no major consistent differences in DF hit rates between the various groups of respondents, whether native or non-native – a pattern which fails to reinforce the notion of any specifically European convergence with regard to DFs.

Since the accents in the stimuli were derived from various Eu-NNS groups, we have also broken down the DF hit rate according to these various groups, as illustrated in Figure 2, where the speakers' native language (i.e. the accent in English stimulus) is shown on the horizontal axis and the judges' native language is shown with the labels and lines.

Figure 2 shows some divergence between the judgements of different groups of Eu-NNS. Dutch, Croatian and Polish listeners selected a proportion of DFs as errors that did not statistically differ from the proportion selected by the NSs (indicated by a dashed line). However, there was a noticeable tendency for the Greek and Spanish participants to report significantly fewer DF realisations as errors than did the NS judges (i.e. below the dashed line). Not only do such patterns raise interesting questions about the rating behaviour of specific groups, but they also suggest that Eu-NNSs are not necessarily more in agreement about their appreciation of other European Englishes than are other groups of judges.

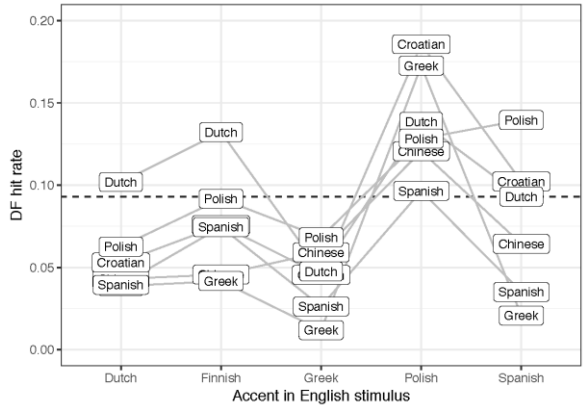

Figure 2. Observed *DF hit rate*, broken down by the speakers' native language (horizontal axis) and the judges' native language (Chinese, Croatian, Greek, Dutch, Polish and Spanish). The observed mean DF hit rate given by native English judges for all speakers is used as a baseline level (indicated by a dashed line).

We have also broken down the DF hit rates according to the type of DF (*fortis* /θ/ versus *lenis* /ð/), as illustrated in Figure 3, where the vertical axis still shows the probability of marking a DF as an error, and where the DF type is shown on the horizontal axis. The listeners' language backgrounds are shown in different colours. For this analysis, the words *with* and *the* (in Sentence 2) were left out of consideration for specific reasons. While the pronunciation of the DF in *with* is subject to regional variation, *the* was considered to be a function word with low perceptual salience, which was not selected even once by any of the respondents in our data set.

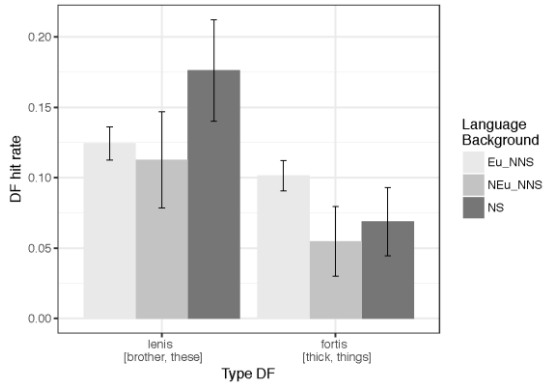

Figure 3. Observed *DF hit rate*, broken down by the type of dental fricatives (fortis or lenis) (horizontal axis) and the judges' language backgrounds. The error bars represent the 95% confidence interval.

Figure 3 shows that overall fortis /θ/ (in the rightmost part) was marked significantly less often as an error than lenis /ð/. Moreover, the NS in our experiment tended to select NNS realisations of /ð/ even more often than did NNSs. This difference in DF hit rate between the NS on the one hand and the Eu-NNS and the NEu-NNS (Chinese) judges on the other is significant. Conversely, the Eu-NNS were inclined to report /θ/ slightly more often as an error than did the NS and NEu-NNS judges.

3.2. Prioritisation

The relative hit rate, which reflects the likelihood of a marked error being a DF, serves as an indication of the prioritisation given to the selection of DFs as opposed to other selected errors. These binary data (DF/other sounds) were also analysed by means of mixed-effects logistic regression (GLMM). The fixed and random factors mentioned in Section 3.1 were also taken into account, and the NSs of English were once again used as the baseline group. The mixed-effects logistic regression analysis of the relative hit rate has been summarised in Table 2.

Table 2. Estimated coefficients for log odds of the *relative hit rate* (i.e. the probability of a marked error being a DF), broken down by respondents' language backgrounds. For fixed effects, regression coefficients are given; for random effects, their standard deviations are given.

Fixed effects	Coefficient	(SE)	Z	P
NS (baseline)	−1.687	(0.557)	−3.031	.002
Eu-NNSs	−0.286	(0.130)	−2.205	.027
NEu-NNSs	−0.385	(0.189)	−2.031	.042
Random effects	Std. Dev.	N		
Judges	0.508	350		
Speakers	0.913	10		
Sentences	0.796	3		
Residuals		11207		

The coefficients reported show that the estimated log-odds for the baseline group of judges, the self-declared NSs of English, was −1.687, corresponding to an observed relative hit rate of 0.213 (thus 21.3% of the marked errors being DFs). The probability of an error being a DF for the Eu-NNS judges (estimated log odds −1.687−0.286, observed relative hit rate 0.166) was slightly lower than for the NSs of English. The NEu-NNS (Chinese) judges selected fewer DFs than other sounds (estimated log odds −1.687−0.385, observed hit rate 0.149), which significantly differed from the NSs. The NSs thus tended to select proportionately more DF errors than other sounds, relative to the Continental European and Chinese judges.

The relative hit rate was also broken down according to the various groups of speakers and judges, as illustrated in Figure 4, where the vertical axis now shows the probability of a DF being marked as an error. The speakers' native language (i.e. the accent in the English stimuli) is shown on the horizontal axis and the judges' native language is represented by means of labels and lines.

Figure 4 reveals that there were some interesting patterns of divergence in the prioritisation of DFs for the different speakers. In particular, the relative hit rate is lower in the Greek and Spanish samples (for all listeners below the dashed line), especially where the judges are NSs of Greek and Spanish themselves. This means that – whenever an error is marked in these samples – it is only rarely a DF. Figure 4 also shows that there was some divergence between the different groups of Eu-NNSs. Unlike the Polish and Croatian judges, Dutch, Greek and Spanish respondents appeared to select DFs significantly less often as errors (as compared to other sounds) than did the NSs of English (indicated by a dashed line). This is evident from their lower relative hit rate. Once again, these patterns suggest that Eu-NNS listeners are not necessarily in agreement about the prioritisation of DFs when judging European Englishes. Some of these tended to agree with NS judges, and others with the Chinese.

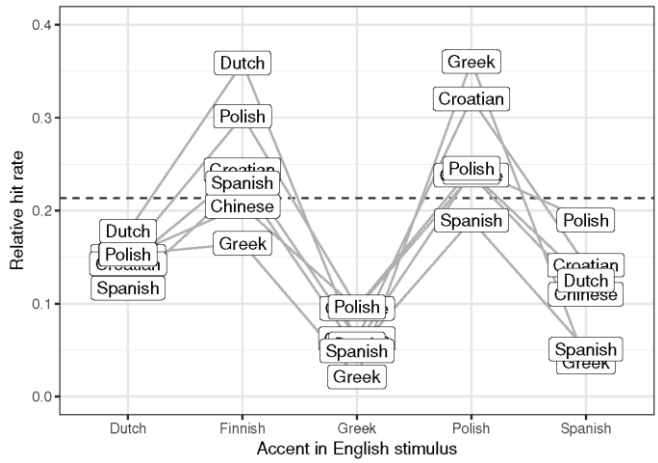

Figure 4. Observed *relative hit rate*, broken down by the speakers' native language (horizontal axis) and the judges' native language (Chinese, Croatian, Greek, Dutch, Polish and Spanish). For each native language, the observed mean relative hit rate given by native English judges for all speakers is used as a baseline level (indicated by a dashed line).

3.3. Evaluation

Error severity was measured on a 5-point scale. Subsequently, this measure was recoded in reverse order, as a result of which higher values signify a more severe judgement of the error (1=least severe, 5=most severe). The average error severity

for all other sounds (observed mean=3.41) was used as a baseline to centralise the data and obtain the relative error severity. This made it possible to determine if any groups evaluated DF errors significantly more or less strictly than other sounds. These relative error severity ratings were analysed by means of mixed-effects linear regression (LMM). Once again, the fixed and random factors mentioned in Section 3.1 were also taken into account, and the NSs of English were employed as the baseline group. The mixed-effects regression analysis of the relative error severity ratings has been summarised in Table 3.

Table 3. Estimated coefficients for the relative error severity (in 5-point scale units), broken down by respondents' language backgrounds. For fixed effects, regression coefficients are given; for random effects, their standard deviations are given.

Fixed effects	Estimate	(SE)	t	
NS (baseline)	0.103	(0.166)	0.621	n.s.
Eu-NNSs	−0.135	(0.153)	−0.880	n.s.
NEu-NNSs	−0.153	(0.223)	−0.686	n.s.
Random effects	Std. Dev.	N		
Judges	0.809	329		
Speakers	0.197	10		
Sentences	0.089	3		
Residuals	0.738	1896		

The coefficients reported indicate that the baseline group of judges, the self-declared NSs of English, showed an observed mean DF error severity (of 0.103) that did not significantly differ from zero (the baseline severity rating for all other sound errors). The Eu-NNSs and the NEu-NNS (i.e. Chinese) judges were as strict as the NSs, as is shown by the non-significant differences. Regardless of language background, all judges rated DFs as severely as the other errors.

A closer inspection of the DF severity estimates and the relative severity estimates did not reveal any significant differences between types of DF assessed (fortis vs. lenis) or between European listener groups. Clearly, while there were significant differences in the identification and prioritisation of DFs, this was not in any way reflected in their evaluation.

4. Discussion and Conclusion

Based on our analysis of patterns of identification, prioritisation and evaluation of non-native realisations of DFs, we were unable to verify any normative convergence among our Eu-NNS respondents. Most importantly, there were no significant differences between the Eu-NNSs and the NS judges in terms of DF hit rates or severity ratings, implying that these groups report roughly the same proportion of DFs as errors and that both groups evaluate these errors equally

strictly. Admittedly, the NSs tended to select proportionately more DF errors than other sounds relative to the Eu-NNSs as a group (the relative hit rate), but in this respect they were actually quite comparable to the Polish and Croatian judges. The Spanish, Greek and Dutch judges, however, reported relatively fewer DF errors – although in the Dutch case, this could be ascribed to their *overall* tendency to report more errors than NSs (see Van den Doel and Quené 2013). In short, there was no indication of any overall consistent differentiation between NSs on the one hand and Eu-NNSs on the other. It is especially striking that there should be no significant differences in evaluation between all groups. This may point to a continuing exonormative orientation towards English /θ/ and /ð/ on the European continent, rather than to any specific convergence between Eu-NNSs. This result is in keeping with other studies which failed to establish any clear evidence of emerging European norms for English, such as Mollin (2006) and Van den Doel and Quené (2013).

Secondly, the Chinese respondents showed lower DF hit rates than did either the NSs or the Eu-NNSs (a difference that was marginally significant), and a correspondingly lower relative hit rate than did the NSs. However, the Chinese judges' severity estimates showed no significant deviation from other groups. In other words, while the Chinese respondents appeared to be less inclined to report any non-native realisations of DFs as errors than the NSs of English, the fewer errors they reported were generally considered to be as severe as they would be by the other two groups. In addition, the Chinese tendency to detect fewer DF errors than some other groups was consistent with their overall low patterns of error detection, a tendency shared with the Greek and Spanish judges. This may suggest that, in terms of normative orientation, subtle differences exist between different groups of NNSs – at least when it comes to highly marked sounds such as /θ/ and /ð/. While the NEu-NNSs and Eu-NNSs may be in agreement with the NSs about the severity of DF errors, the Chinese respondents were less inclined, able or willing to report these as such. It would therefore seem unwarranted to generalise about NNSs as a group – at least where perceptions of DFs are concerned. Similarly, Van den Doel's (2006) investigation of NS evaluation of Dutch-accented pronunciation features showed that listener groups agreed on the gravity of specific cases of **th**-stopping, yet show structurally different patterns of detection (297). Van den Doel concluded from this that NS perceptions of DFs in NNS speech are far from uniform, and this may now be extended to NNSs.

Evidently, our data do not show any evidence of convergence among Continental Europeans, or even among groups of NNSs, in their judgements of DF realisations. Nonetheless, some idiosyncratic patterns present themselves in specific groups of NNS listeners. While the Chinese, Greek and Spanish listeners reported fewer DFs as errors, the Polish and Croatian listeners distinguished themselves by selecting relatively more of these than did other NNSs. These different tendencies may be accounted for in a number of ways, ranging from perceptual difficulties on the one hand to truncated repertoires in error detection and pedagogical traditions on the other – as is, for instance, done in Beinhoff

(2014), which showed an effect of proficiency in NNSs' evaluations of Spanish-accented speech. It would, however, be interesting to consider if transferred L1 norms can be invoked to explain some of the variation between the different groups of NNS listeners.

The notion that DFs are much less of a priority to NNSs because of their conspicuous absence in most phoneme inventories suggests a correlation between NNSs' attitudes to non-native realisations of /θ/ and /ð/ and their inclusion in listeners' first languages. However, we did not find any evidence for such a correlation. Even though there were some specific differences between the Greek and Spanish respondents and the NSs in terms of DF hit rate, these can be explained by pointing at overall group tendencies towards error detection and evaluation. Moreover, the Greek tendency to underreport errors of any kind (including DFs) may be difficult to explain on the basis of their L1. In addition, the lack of significant difference between Dutch, Polish and Croatian judges (whose languages do not feature /θ/ and /ð/), and the NS judges (most of whom are likely to be at least familiar with these sounds) on the other, may even be taken to mean that familiarity with DFs does not necessarily affect listeners' judgements – as was also found in Beinhoff (2008) for Greek and German judges.

It may be speculated that especially in the context of the present experiment, where both NS and NNS listeners were asked to report on specific errors, judges would be more inclined to resort to some sort of external, prescriptive pronunciation model rather than rely on their own judgements or any transferred L1 norms, but this cannot be verified with our data. Arguably, this could affect NNS judges disproportionately, if it is believed that NNSs by their very nature adopt NS or classroom norms more uncritically than do other groups. This would imply that, as "victims of an ideology that is imposed on them" they have been "brainwashed" by so-called native-speakerism (Jenkins 2007: 59, 187). However, if NNSs are viewed as a diverse group of stakeholders who do not respond to the languages they are exposed to as passive, docile and unthinking recipients (cf. Holiday 2006) and who may be credited with as much agency as NSs, this would argue against such conceptualisations of what should perhaps be termed "*non*-nativespeakerism". But even if we accept that any error bias will play less of a part in more realistic tasks or situations, the fact remains that the experimental conditions were identical for all respondents.

The absence of any effects which may be ascribed unequivocally to transferred L1 norms, and the lack of any significant divergence between Eu-NNSs and NSs, suggest that in this instance, local influences play a subordinate role. In addition, it seems unlikely that the similarity between the two groups can be explained by positing the adoption of more accommodating lingua franca norms for DFs by NSs. This is because non-native realisations of /θ/ and /ð/ are among the most commonly reported sources of "error" for all groups. What can be claimed, however, is that the convergence between Eu-NNSs and NSs cannot simply be extended to include the Chinese listeners, whose lower detection rate suggest either a relative unwillingness, or reduced ability, to report non-native realisations

of DFs as errors. Either could be taken as evidence that Chinese judges resort less to exonormative NS standards for the detection, prioritisation and evaluation of these sounds than do other groups of listeners. If so, it could be argued that a Lingua Franca Core which does not include the teaching of /θ/ and /ð/ to NNSs (as posited in Jenkins 2000) may be more relevant to some NNSs than to others.

Admittedly, our analysis shows that NSs were on the whole more inclined than NNSs to prioritise DFs in their identification of potential errors. A closer inspection of the data reveals, however, that this effect is due to the higher frequency with which the NS judges tended to select NNS realisations of /ð/ than did NNSs. Since this did not affect the evaluation of DF errors by any groups, perceptual differences between groups may have played a part in this. While both NS and NNS listeners may readily confuse /θ/ with other sounds (Hanulíkóva and Weber 2012: 615), it should be pointed out that /ð/ is a notoriously weak sound and may have less perceptual salience to NNSs. In any event, our results suggest clearly that it would be inadvisable to generalise about NSs' and NNSs' attitudes to DF errors without considering whether /θ/ and /ð/ errors may be perceived differently by specific groups. Such differentiation is in keeping with Van den Doel (2006: 239), who found that North American judges evaluated /θ/ and /ð/ errors structurally differently from other NS respondents. If it is true that specific groups of judges, whether NS or NNS, attach more importance to /ð/ than /θ/ errors, this would suggest, pace Jenkins (2000), that categorical claims about the relative insignificance of all DF errors to all groups of NNSs will need to be revisited. At the very least, it would be prudent to refrain from making such premature claims in textbooks aimed at any such groups (such as Walker 2010).

In an experiment of this kind, there will necessarily be a number of limitations. For instance, we were unable to take into consideration factors such as respondents' proficiency, which was not measured objectively, or their general attitudes to evaluation and language learning. In fact, as was also pointed out in Van den Doel and Quené (2013), it would have been interesting to explore the effect of "educational traditions which either favour or disfavour ambitious standards for language learning (and pronunciation training in particular)" (pp. 91–92), and of any local attitudes to "strictness and precision in education in general" (Edgar Schneider, as quoted in Van den Doel and Quené 2013). Since we used the same data set as in Van den Doel and Quené (2013), we have not been able to incorporate this.

Surveys such as EF-EPI (n.d.) and the Eurobarometer (2006) are routinely cited to support stereotypical impressions of proficiency in English in different groups of NNS, but these often either rely on very limited data or on self-reporting. Even if more inclusive or objective criteria are used to differentiate between different groups of NNSs in terms of proficiency, this does not help to account for all the findings of the present experiment, such as lack of significant differences between Dutch, Croatian and Polish respondents. It would therefore indeed be a good idea to investigate the link between proficiency and error assignment in any follow-up studies, as in Beinhoff (2014). This would be one way in which the

differential assessment of DFs by different groups of NNSs can then be more explicitly linked to effects in specific groups. This may be more insightful than attributing this to a possibly artificial, or even spurious, distinction between NSs on the one hand and NNSs on the other. In fact, it could be argued that all differences between groups of NNSs should be investigated separately, and that differences in language attitudes, in educational traditions and in perceptual difficulties should be taken into account, before any conclusions can be drawn about NNSs' overall assessment of DFs as a homogeneous group.

It would also be interesting to consider the effects of different DF-substitutions on NNS users of English. This is an issue not addressed in any detail in the LFC or in the present study, but it would seem highly unlikely that all substitutions will be regarded with the same level of leniency, given the part that some of these substitutions play in the recognition and stigmatisation of specific accents, such as **th**-alveolarisation in French and German English. Unfortunately, we do not have sufficient data to report on this at present. If it turns out, however, that such substitutions are evaluated very differently by NNSs, it would emphasise the importance of factors other than "mere intelligibility" in a lingua franca context.

As it is, there is no indication that non-native substitutions of /θ/ and /ð/ are evaluated more leniently by Eu-NNSs than by their NS counterparts. If we accept Jenkins's (2000) position that NNSs do not perceive such substitutions to be harmful to intelligibility, and at the same time find that they continue to report these as errors, we should entertain the possibility that some judges are swayed by considerations of acceptability. It may be suggested, of course, that such normative behaviour is merely reflective of internalised native-speakerism, and does not merit serious consideration other than as commentary on pervasive language teaching ideologies. However, if, as Jenkins has also suggested, "NNSs should have input into the determining of their pronunciation norms" (2007: 26), the views of different NNS stakeholders, whether exonormative, proficiency-based or the product of local normativity, should be reflected in this. This would include allowing for the possibility that some NNSs are more aware than others of the stigmatisation of specific accent features, and may even actively contribute to this. Jenkins (2000: 160) has claimed that "[t]here really is no justification for doggedly persisting in referring to an item as "an error" if the vast majority of the world's L2 English speakers produce and understand it". However, if we are to take the input of NNSs seriously, it may be argued that precisely such a justification may be found in NNSs' observed persistence in reporting DF-substitutions as errors.

References

Bates, Douglas, Maechler, Martin, Bolker, Ben and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67. 1-48.

Beinhoff, Bettina. 2008. Looking for the 'real' native speaker: the perception of native and non-native English accents by non-native speakers of English. In Ewa Waniek-Klimczak (ed.), *Issues in Accents of English*, 120-139. Newcastle: Cambridge Scholars Publishing.

Beinhoff, Bettina. 2013. *Perceiving identity through accent: attitudes towards non-native speakers and their accents in English.* Oxford: Peter Lang.

Beinhoff, Bettina. 2014. Perceiving intelligibility and accentedness in non-native speech: a look at proficiency levels. *Concordia Working Papers in Applied Linguistics* 5. 58-72.

Blommaert, Jan. 2010. *The sociolinguistics of globalization.* Cambridge: Cambridge University Press.

Breiteneder, Angelika. 2009. English as a Lingua Franca in Europe: an empirical perspective. *World Englishes* 28. 256-69.

Buschfeld, Sarah. 2011. *The English language in Cyprus: an empirical investigation of variety status.* PhD dissertation University of Cologne, Germany.

Collins, Beverley and Inger M. Mees. 1996. *The PHONETICS of English and Dutch.* Leiden: Brill.

Crystal, David. 2001. The future of Englishes. In Anne Burns and Caroline Coffin (eds.), *Analysing English in a global context*, 53-64. London: Routledge.

Derwing, Tracey and Murray Munro. 2015. *Pronunciation fundamentals: evidence-based perspectives for L2 teaching and research.* Amsterdam: John Benjamins.

Devonish, Hubert and Otelemate Harry. 2004. Jamaican Creole and Jamaican English: phonology. In Edgar Schneider, Kate Burridge, Bernd Kortmann, Rajend Mesthrie and Clive Upton (eds.), *A handbook of varieties of English, vol. 1*, 450-480. Berlin: Mouton.

Edwards, Alison. 2016. *English in the Netherlands: functions, forms and attitudes.* Amsterdam: John Benjamins.

Eurobarometer. 2006. Special eurobarometer 64.3: Europeans and their languages. [Online] Available from http://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/-ebs_243_en.pdf. [Accessed: 20 July 2018].

EF-EPI. n.d. The world's largest ranking of countries by English skills. [Online] Available from http://www.ef.nl/epi/. [Accessed: 20 July 2018].

Forche, Christian. 2012. On the emergence of Euro-English as a potential European variety of English: attitudes and interpretations. *Jezikoslovlje* 13. 447-478.

Hanulíkóva, Adriana and Andrea Weber. 2012. Sink positive: Linguistic experience with the substitutions influences nonnative word recognition. *Attention, Perception, and Psychophysics* 74. 613-629.

Holiday, Adrian. 2006. Native-speakerism. *ELT Journal* 60. 385-387.

Hosmer, David W. and Stanley Lemeshow. 2000. *Applied logistic regression.* New York: Wiley.

Hickey, Raymond. 2004. *Irish English: history and present-day forms.* Cambridge: Cambridge University Press.

Jenkins, Jennifer. 2000. *The phonology of English as an international language: new models, new norms, new goals.* Oxford: Oxford University Press.

Jenkins, Jennifer. 2001. 'Euro-English' accents. *English Today* 68. 16-9.

Jenkins, Jennifer. 2007. *English as a lingua franca: attitude and identity.* Oxford: Oxford University Press.

Jenkins, Jennifer. 2017. An ELF perspective on English in the post-Brexit EU. *World Englishes* 36. 343-346.

Jenkins, Jennifer, Cogo, Alessia and Martin Dewey. 2011. Review of developments in research into English as a Lingua Franca. *Language Teaching* 44. 281-315.

Kautzsch, Alexander. 2014. English in Germany: Retreating exonormative orientation and incipient nativization. In Sarah Buschfeld, Thomas Hoffmann, Magnus Huber and Alexander Kautzsch (eds.), *The Evolution of Englishes: The dynamic model and beyond*, 203-228. Amsterdam: John Benjamins.

Kirkpatrick, Andy. 2007. *World Englishes: implications for international communication and English language teaching*. Cambridge: Cambridge University Press.

Klimczak-Pawlak, Agata. 2014. *Towards the pragmatic core of English for European communication.* Cham: Springer.

Kreft, Ita and Jan de Leeuw. 1998. *Introducing multi-level modeling*. London: Sage Publications.

Ladefoged, Peter and Ian Maddieson. 1997. *The sounds of the world's languages.* Oxford: Blackwell.

Luke, Douglas. 2004. *Multilevel modeling*. Thousand Oaks, Ca.: Sage.

MacKenzie, Ian. 2014. *English as a lingua franca: theorizing and teaching English.* Abingdon: Routledge.

MacKenzie, Ian. n.d. The allophonic rules of Spanish. [Online] Available from: https://www.staff.ncl.ac.uk/i.e.mackenzie/allophonr.htm. [Accessed: 20th July 2018].

Modiano, Marko. 2017. English in a post-Brexit European Union. *World Englishes* 36. 313-327.

Mollin, Sandra. 2006. *Euro-English: assessing variety status.* Tübingen: Narr.

Pampel, Fred. 2000. *Logistic regression: a primer*. Thousand Oaks, CA: Sage Publications.

Quené, Hugo and Huub van den Bergh (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language* 59. 413-425.

R Development Core Team. 2016. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. [Online] Available from http://www.R-project.org/ [Accessed: 1 December 2016].

Salakhyan, Elena. 2012. The emergence of Eastern European English. *World Englishes* 31: 331-350.

Schneider, Edgar. 2004. Global synopsis: phonetic and phonological variation. In Edgar Schneider, Kate Burridge, Bernd Kortmann, Rajend Mesthrie and Clive Upton (eds.), A *handbook of varieties of English, vol. 1*, 1111-1138. Berlin: Mouton.

Schneider, Edgar. 2017. The linguistic consequences of Brexit? No reason to get excited! *World Englishes* 36. 353-355.

Seidlhofer, Barbara. 2001. Towards making "Euro-English" a linguistic reality. *English Today* 68. 14-16.

Van den Doel, Rias. 2006. *How friendly are the natives? An evaluation of native-speaker judgments of foreign-accented British and American English*. Utrecht: LOT.

Van den Doel, Rias. 2008. The blind spots of Jenkins's Lingua Franca. In Ewa Waniek-Klimczak (ed.), *Issues in Accents of English*, 140-149. Newcastle: Cambridge Scholars Publishing.

Van den Doel, Rias, and Hugo Quené. n.d. The accents of European English: an online experiment. [Online] Available from let.uu.nl/~Rias.vandenDoel/personal/wwstim/eureng/html/. [Accessed: 20 July 2018].

Van den Doel, Rias, and Hugo Quené. 2013. The endonormative standards of European English: emerging or elusive? *English World-Wide* 34. 77-98.

Walker, Robin. 2010. *Teaching the pronunciation of English as a lingua franca.* Oxford: Oxford University Press.

Weinberger, Steven. 2011. Speech accent archive. [Online] Available from http://accent.gmu.edu. [Accessed: 20 July 2018].

Wells, John. 2008. With. *Longman pronunciation dictionary*. 3rd edn. Edinburgh: Pearson, 904.

Wells, John. 2010. Exotic dental fricatives. John Wells' Phonetic Blog. 7 April 2010. [Online] Available from http://phonetic-blog.blogspot.nl/2010/04/exotic-dental-fricatives.html. [Accessed: 20 July 2018].

PARENTHETICAL CLAUSES IN THE QUR'ĀN

YEHUDIT DROR
University of Haifa, Israel
judror@gmail.com

Abstract
In the few traditional Arabic grammatical sources that address the term parentheticals it is usually defined as the insertion of a clause between two other clauses, or between two syntactic components, for ta'kīd "emphasis." In this article I examine Qur'ānic parenthetical clauses in the theoretical framework of relevance theory. It transpires that the parenthetical clause is placed where it achieves optimal relevance and therfore the conveyed utterance does not require the addresses to waste any efforts trying to procees the information and correctly interpret it. Optimal relevance also means having a contextual effect. The Qur'ānic parenthetical clauses have one of the following contextual effects: They serve to affirm God's omnipotence, indicating that only God produces suras, created heaven and earth. He is the forgiver and all depends on His will; to explain what it meant by a specific statement or to explain the reason behind a certain action; to qualify, to highlight a specific characterization, for example, one of the parenthetical clauses modify the Qur'ān as the truth from God; to provide background information, which could explain further developments in the narrative.

Keywords: parenthetical clauses, Relevance Theory, contextual effect

1. Introduction

1.1. General definition of the term "parentheticals"

This article investigates parenthetical clauses in the Qur'ān, a category which subsumes a wide range of forms and lacks a clear definition in either traditional Arab grammar or Western research on Arabic language. As a general characterization of parentheticals, the term *parenthesis* may be said to refer to any peripheral information, expressed through a single word, phrase or clause, and in terms of content the information is external to the sentence (Zewi 2007: 2).[1] Parentheticals are elements varying in length and complexity, category and function, as in:

a) One-word parentheticals: "Umm I don't have a sleeping bag *unfortunately*."
b) Nominal apposition: "She claimed that the new Prime Minister *Jim Callaghan* had offered his predecessor the job of Foreign Secretary in his government."

[1] Cf. Ziv (1985: 181).

c) Question tags: "He suffered great mental distress *didn't he* after the war?"
d) Clauses: "It's been a mixture of extreme pleasure *I've had hundreds of letters from all sorts of people who have enjoyed the book* and considerable irritation because of being constantly interviewed" (Dehé and Kavalova 2007: 2–3).[2]

As seen in these examples, parentheticals are linearly[3] integrated in the host sentence, i.e., they intersect with other structures in it on a linear plane, sharing with them a terminal string; nor are they are linked to the host by any syntactic nodes (Kaltenböck 2007: 26)[4] therefore they seem to be independent. It is also claimed that they interrupt the prosodic flow of the sentence by introducing intonational breaks in the host sentence (Dehé and Kavalova 2007: 1). Parentheticals are also characterized by the lack of a specified position in the sentence. They may occur in initial, medial or final position in the sentence. However, there are some restrictions regarding their position in it: for example, the following sentences are grammatically correct: "*As far as I know*, Bill sent her a lot of money"; "Bill sent her a lot of money, *as far as I know*"; "Bill, *as far as I know*, sent her a lot of money." However, a sentence like *Bill sent her, as far as I know, a lot of money" is ill-formed because it breaks an inseparable sequence; in other words, there are weak spots in the syntax of the host sentence which enable the insertion of a parenthetical expression more readily than others (Ziv 1985: 182).[5] For example, inserting a parenthetical between a subject and its verbal predicate is acceptable, while inserting it between a verb and its direct object is regarded as ill-formed (Peterson 1999: 239).

As for the syntactic relations between the parenthetical clause and its host sentence, some scholars[6] show that they do exist. For example, anaphors in a parenthetical can be bound by antecedents in the host clause. Other scholars, such as Peterson (1999: 230, 232),[7] argue that parentheticals are non-syntagmatic – i.e, they involve a lack of linkage between the elements and therefore do not constitute a grammatical construction.

The syntactic independence of the parenthetical from its host sentence can be demonstrated by a number of tests; I mention only a few:

(a) They can be deleted, e.g., "John talked to us, *it seems*, about literature and Mary did too." The intended meaning is that Mary talked about literature

---

2 For additional examples of parentheticals in English, see Kaltenböck (2007: 29–30).

3 Burton-Roberts (2006: 180) explains the term *linearity* as follows: "In syntactic order, linear order is generally held to be a function of hierarchical syntax structure: order is determined by, and within, constituent domain So, if one expression is contained by another expression on the linear axis, it should be contained by the expression on the hierarchical axis. In other words, it should be a syntactic constituent of that expression."

4 Cf. Dehé and Kavalova (2007: 1); Kavalova (2007: 145).

5 Cf. Kaltenböck (2007:42–43).

6 For example, Kavalova (2007) refers to Hoffmann (1998), Jackendoff (1977) and Potts (2002).

7 Haegeman (1988) is also mentioned in the literature as one of the scholars who support the unintegrated approach of parentheticals.

and not that Mary talked, it seems, about literature. The parenthetical does not form a component with the VP (talked) of the host.

(b) Parentheticals cannot be the focus of a cleft sentence, e.g., "Emmanuel Shinwell thought – *and he is after all a previous Defence Minister* – that you can't have informed opinion on this vital mater without…." It is impossible to construct this sentence as *It was (and) *he is after all a previous Defence Minister* who/that Emanuel Shinwell thought that you can't have….

(c) Parentheticals are not temporally subordinated to the host, e.g., "In fact it was very candidly told *and I repeat the acknowledgment of the candour* with which it was placed before him in January." The present tense in the parenthetical is independent from the past tense in the host (Kaltenböck 2007: 35).[8]

Finally, pragmatically parentheticals express a comment by the speaker or they may strengthen or weaken its force, or specify the form of the speaker's attitude to the content of the utterance (Ziv 1985: 182).

As for the definition of the term *parenthetical* in medieval Arabic grammatical treaties, we might refer to Ibn Ǧinnī (1913: vol. 1,339). In a chapter entitled *bāb fī l-iʻtirāḍ* "Chapter on parenthesis" [*lit*. "Chapter on the interception"] he says that it is a known feature which occurs in the Qur'ān, in poetry and in prose. Its main function is *ta'kīd* "emphasis," therefore grammarians, rhetoricians, poets and writers integrate in their writings sentences, where a syntactic constructor is interpolated between subject and predicate or between other components which may not be separated.

Ibn Hišām (1991: vol. 2, 21) refers to the term *ǧumla muʻtariḍa* "parentheticals clause," classifying it as one of the clauses devoid of syntactic status (*ǧumlatun lā maḥalla lahā mina l-'iʻrābi*)[9]. He explains this type of clause as follows:

> al-muʻtariḍatun bayna šay'ayni li-'ifādati l-kalāmi taqwiyatan wa-tasdīdan 'aw taḥsīnan, wa-qad waqaʻat fī mawāḍiʻa
> "The parenthetical clause [is inserted] between two things [i.e., between two clauses or between two syntactic constituents] to enhance the speech by reinforcing it, by focusing [on a certain meaning], or by embellishing it and it stands in [various] positions."[10]

---

8 Cf. Kavalova (2007: 158–160).

9 The term *ǧumlatun lā maḥalla lahā mina l-'iʻrābi* means that it has no specified syntactic function in the sentence that contains it. This clause does not function as head, complement or adjunct.

10 Gully (1995: 78) explains Ibn Hišām's definition as follows: "The function of parenthesis is the separation of two clauses by another clause; in essence, the reinforcement and sealing off, or embellishment of speech in a manner which actually enhances the meaning."
Reckendorf (1921: 382) also explains that parenthetical clauses (*ǧumla muʻtariḍa*) stand in the middle, i.e., between two clauses.

According to these definitions the basic criterion for identifying parenthetical clauses is that they are inserted between two clauses or two syntactic constituents. Both Ibn Ǧinnī and Ibn Hišām further indicate the most typical places of interpolation for parenthetical clauses, of which I will mention a few cases:

(a) Between subject and predicate: *wa-qad ʾadrakatnī* ***wa-l-ḥawādiṯu ǧammatun*** *ʾasinnatu qawmin lā ḍiʿāfin wa-lā ʿuzli* "The spears [belonging] to the people, who are neither weak nor defenseless, have reached me, while the events are numerous" [*lit*. I was reached by — while the events are numerous — the leaders of the people, who are neither weak nor defenseless]" (Ibn Ǧinnī 1913: vol. 1, 340).[11]

(b) Between *protasis* (the clause containing the condition) and *apodosis* (the clause containing the *conclusion)*: *ʾin yakun ġaniyyan ʾaw faqīran* ***fa-llāhu ʾawlā bihimā*** *fa-lā tattabiʿū l-hawā* (Q 4: 135) "[O believers, be upholders of justice, witnesses for God, even though it be against yourselves, or your parents or near relatives], whether rich or poor **God stands close to them, so** follow not [your] lust." (Ibn Hišām 1991: vol. 2, 27).[12]

(c) Between a noun and its adjective: *wa-ʾinnahu la-qasamun* ***law taʿlamūna*** *ʿaẓīmun* (Q 56: 76) "[I swear by the place of the stars] and that is a mighty oath, if you but know" [*lit*. and that is an oath, **if you but know**, mighty] (Ibn Ǧinnī 1913: vol. 1, 339).[13]

(d) Between a preposition and its governed noun: *ištaraytuhu bi****ʾarā*** *ʾalfi dirhamin*
"I bought it, **I think**, for a thousand dirhams" (Ibn Hišām 1991: vol. 2, 30).

(e) Between two clauses: *fa-ʾtūhunna min ḥayṯu ʾamarakumu llāhu* ***ʾinna llāha yuḥibbu t-tawwābīna wa-yuḥibbu l-mutaṭahhirīna*** *nisāʾukum ḥarṯun lakum* (Q 2: 222–223) "[**They ask you about menstruation. Say: It is painful; so avoid (intercourse with) women during menstruation and do not approach them till they are clean. When they have cleaned themselves,] then you may go unto them as Allah has commanded you. Verily, Allah loves those who repent constantly, and (He) loves those who purify themselves. Your women are a tillage for you**" (Ibn Hišām 1991: vol. 2, 32).

Also, Ibn Hišām's definition implies a connection between the host sentence and the parenthetical clause because it establishes some pragmatic link between them. The parenthetical clauses may reinforce, affirm or specify what has been uttered previously in the sentence.

---

11 Cf. Ibn Hišām (1991: vol. 2, 21).

12 The translation of the Qurʾānic verses is taken from Arberry (1964), while I made a few changes in the original translation.

13 Cf. Ibn Hišām (1991: vol. 2, 28).

1.2. Objectives and frame of work

Despite Arab grammarians' occasional references to the term ǧumla mu'tariḍa,14 this notion has not been explicitly defined – a fact that may explain three difficulties that arose during the research:

(a) To create a comprehensive list of parenthetical clauses I referred to the exegetical literature while trying to detect all occurrences of the term *ǧumla mu'tariḍa*. It seems that the lack of criteria for syntactically identifying parentheticals has led to various options among the commentators, and in many cases one commentator identifies varied constructions of parentheticals which are not mentioned by the others.

(b) Another prominent problem in the exegetical literature is the complexity of determining the exact border between parenthetical units and clauses such as circumstantial clauses, e.g., *yurīdūna 'an yaḫruǧū mina n-nāri* ***wa-mā hum bi-ḫāriǧīna minhā*** *wa-lahum 'aḏābun muqīmun* (Q 5:37) "They will desire to come forth from the Fire, **but they will not come forth from it**, for them awaits a lasting chastisement." According to Šawkānī (1997: vol. 2, 55), the clause in bold might be analyzed as a circumstantial clause or as a parenthetical clause. Ibn Hišām (1991: vol. 2, 443) distinguishes these two types, saying *inter alia* that a parenthetical can be an interrogative clause or a clause expressing a command. A circumstantial clause, on the other hand, must be informative. However, this cannot be regarded as a clear-cut rule for distinguishing the two types because in various cases such as Q 5: 37 or 2: 83, there is an informative clause which commentators explain as both as parenthetical and circumstantial.

(c) The third problem concerns the relation between the position of the parenthetical clause and its pragmatics, as the following example shows: *wa-llaḏīna 'iḏā fa'alū fāḥišatan 'aw ẓalamū 'anfusahum ḏakarū llāha fa-staġfarū li-ḏunūbihim* ***wa-man yaġfiru ḏ-ḏunūba 'illā llāhu*** *wa-lam yuṣirrū 'alā mā fa'alū wa-hum ya'lamūna* (Q 3: 135) **"And those who, when they commit an indecency or do injustice to themselves, remember God and seek forgiveness for their sins — and who forgives the sins except God? — And they do not knowingly persist in what they have committed."** According to Ḥasan (2012: 84), the motivation for introducing the parenthetical clause into the sentence is to emphasize the idea that no one except God can forgive peoples' sins. I find that such an explanation is not sufficiently convincing and fails to explain the reason for the insertion. The idea of God as forgiver could

14 References to the term *al-i'tirāḍ* are found in Arabic rhetoric. For example, 'Askarī (1971: 410) defines it as interruption of the host sentence. He explains that the insertion breaks the flow, hence there is structural incompleteness of the host. However, after this break the listener returns to the host sentence, which becomes a complete sentence semantically and syntactically.

equally have been uttered if the parenthetical clause had been in final-position.

This article contributes to the discussion of the relation between the host sentence and parentheticals in the Qur'ān. It has two goals:

1) To explain in what way parenthetical clauses are linked to their host clauses and whether the various types of parentheticals differ in this respect. This issue is addressed in section 2.
2) To draw attention to the reason for introducing a parenthetical clause into the main utterance. Using the framework of Relevance Theory (Sperber and Wilson 1996), in section 3 I explain what governs the use and the position of such structures in the Qur'ān.

Additionally, this study might contribute to the understanding of the relationship between phrases, clauses, and sentences in the Qur'ān. I refer to the relevance theory as a way which attempts to explain how parentheticals are related to the host sentence and how they contribute to the interpretation of the speaker's intention. However, I should mention in this context the pioneering work of Salwa El-Awa on textual relations in the Qur'ān. In this work, El-Awa also discuss the problem of textual relations in the Qur'ān according to the principles of the Relevance Theory.[15]

The methodology employed in the current study and in El-Awa's work might encourage future Qur'ānic studies dealing with word order, information structure and related issues.

A source which already identified parenthetical clauses in the Qur'ān is the exegetical literature. I used the search engine in the Internet site http://www.altafsir.com/index to collect all occurrences of parenthetical clauses. Out of 127 parenthetical clauses, only 22 are discussed here because I ruled out three types of clauses: first, clauses that can be analyzed as both circumstantial and parenthetical. For example, *wa-llāhu yaḥkumu* ***lā mu'aqqiba li-ḥukmihī*** (Q 13: 41) "And God judges, **no one repels His judgment**" is usually analyzed as a circumstantial clause; only a few commentators raise the option that it might be a parenthetical clause.[16] Second, clauses starting with the anaphoric pronoun *hāḏā* "this" or *ḏālika* "that" or *'ulā'ika* "those," for example, Q 65: 1. I could not understand what differentiated these clauses from other clauses of this type which are not analyzed as parentheticals. Third, sometimes I identified another function of the clause other than parenthetical. For example, in numerous cases the commentators regard a conjuncted clause as parenthetical, for example, Q 2: 200 (*fa-min n-nāsi*…). However, my examination showed that these clauses can be analyzed as coordinated clauses because they usually refer to a new subject. Parenthetical clauses, on the other hand, refer to the same subject mentioned previously in the host sentence.

---

15 See: El-Awa (2006: 1–2).

16 See: Šawkāni (1997: vol. 3, 125)

2. Two Types of Parenthetical Clauses

We can distinguish two types of parenthetical clauses in the Qur'ān: middle-positioned and final-positioned.[17] Each type has the following properties:[18]

2.1. Syntactic relation vs. non-syntactic relation

(1)
wa-man 'aḥsanu dīnan mimman 'aslama wağhahu li-llāhi wa-huwa muḥsinun wa-ttaba'a millata 'ibrāhīma ḥanīfan **wa-ttaḫaḏa llāhu 'ibrāhīma ḫalīlan** (Q 4: 125)
"And who is better in religion than he who submits his will to God, he being righteous and one who follows the creed of Abraham, a man of pure faith? **God took Abraham for a friend**."

Syntactically, the parenthetical clause seems connected to the host by the coordinating conjunction *wa-* "and"; however, the particle *wa-* functions here as *wāw al-isti'nāf*[19] "*wāw* of commencement," which indicates the beginning of a new sentence wholly independent from the sentence that precedes it. According to the commentators, the communicative value of the parenthetical clause is to emphasize the obligation to follow the creed of Abraham by stating that he is chosen as a friend of God. Had this clause functioned as *ğumla ma'ṭūfa* "coordinated clause" connected to the previous clause, it would have been meaningless (*law ğa'altahā ma'ṭūfatan 'alā l-ğumlati qablahā lam yakun lahā ma'nā*) (Zamaḫšarī 1947: vol. 1, 569).[20] It means that coordination of two or more clauses might indicate that the speaker, having mentioned one thing, wants to go on to talk about something else that is similar, but this is not the case in Q 4: 125. The parenthetical clause adds some information about a previous subject (Abraham).

While there is no syntactic relation between the parenthetical clause and the host, there is a semantic bond between the two which is achieved by the fact that one component (the name of Abraham) is shared by the host sentence and the parenthetical clause. The repetition of this name was probably intended to prevent any ambiguity as to the identity of the person who is considered God's friend. Had the proper name been replaced by a pronoun ****wa-ttaḫaḏahu llāhu ḫalīlan*** "God took him as a friend" the suffixed pronoun *-hu* could have had two possible antecedents: Abraham, or he who submits his will to God.

---

17 The occurrence of parenthetical clauses in final-position is mentioned in Arabic rhetoric, where they are also called *al-i'tirāḍ at-taḏyīlī* "suffixed parentheticals" and "final-positioned parentheticals." For example, Qazwīnī (2002: 129) says that this type of parentheticals is introduced to complete the utterance.

18 Examples (8) and (9) are middle-positioned parentheticals even though they have the same properties as the final-positioned parentheticals – namely, there is no syntactic relation between the parenthetical clause and the clause which precedes it.

19 See: Yāqūt (1998: vol. 4, 288).

20 Cf. Nasafī (1996: vol. 1, 368).

In the following example, the connection between the host and the parenthetical is licensed by grammar. It is inserted between the subject and the predicate and it includes a pronoun-antecedent relation, where the suffixed pronouns *-him* (in *ʻalayhim*) and *-hum* (in *ʼanḏartahum*, *tunḏirhum*) refer back to the unbelievers:

(2)
ʼinna llaḏīna kafarū sawāʼun ʻalayhim ʼa-ʼanḏartahum ʼam lam tunḏirhum lā yuʼminūna (Q 2:6)
"The unbelievers, it is all the same to them whether you have warned them or have warned them not; they do not believe."

2.2. Dependent clauses vs. independent clauses

Full interpretation of the parenthetical clauses, which are syntactically connected to the host in an anaphoric relation, can be achieved only in a given context, namely when they are integrated in the host sentence. On the other hand, parenthetical clauses without any syntactic relation are independent and *self-contained* clauses. This argument can be tested by taking the parenthetical clause out of its context:

(3)
fa-ltaqaṭahu ʼālu firʻawna li-yakūna lahum ʻaduwwan wa-ḥazanan ***ʼinna firʻawna wa-hāmāna wa-ǧunūdahumā kānū ḫāṭiʼīna*** (28:8)
"Then the people of Pharaoh picked him out that he might be unto them an enemy and sorrow to them; verily Pharaoh and Haman [Pharaoh's vizier] and their hosts were sinners."

Q 28:8 belongs to a larger thematic unit which is the narrative of Moses (verses 1–44).

Verses 4–5 relate that **Pharaoh exalted himself in the land (of Egypt) and divided its people into sections, weakening a group of them; he slaughtered their sons and spared their females.** However, God **intended to bestow His favor upon those who were oppressed on earth by giving them power on earth, and to show Pharaoh and Haman what they dreaded from them**. In light of this information, we read in verse seven that Moses' mother is requested to breastfeed Moses and she is told that **when she feels fear for Moses she should cast him into the river without any hesitation because he will return as one of the apostles.** When we read verse eight we understand that Moses' mother indeed cast him away and now he is picked up by Pharaoh's people. This verse ends with the parenthetical clause, while no syntactic relation exists between it and the previous sentence. However, the semantic relation is clear: the parenthetical clause explains the reason for sending Moses to Pharaoh in order to cause them sorrow: both Pharaoh and Haman were sinners and they must be punished. However, I argue that even when the parenthetical is used outside its context it is a comprehensive utterance. **The clause "Verily Pharaoh and Haman and their**

hosts were sinners" is a statement which can stand by itself because the grammar and the semantics of the parenthetical clause in no way depend on the host. In example (4) we see the opposite case. The syntactic relation indicates that the parenthetical clause is a dependent utterance that must be linked to the host to be fully interpreted.

(4)
wa-la-'in 'aṣābakum faḍlun mina llāhi la-yaqūlanna ***ka-'an lam takun baynakum wa-baynahu mawaddatun*** *yā-laytanī kuntu ma'ahum fa-'afūza fawzan 'aẓīman* (4: 73)
"**But if a bounty from God befall you, he will surely say – as if there had not been any affection between you and him – If only I had been with them I would have achieved a mighty triumph!"**

The parenthetical clause *ka-'an lam takun baynakum wa-baynahu mawaddatun* cannot be taken out of its context because the anaphors in the parenthetical clause (*-kum baynakum* and *-hu* in *baynahu*) are bound by antecedents in the host sentence. Furthermore, unlike example (3), the parenthetical clause in example (4) can be characterized as the speaker's comment, a term which is explained in the following section.

2.3. Subject-oriented parentheticals vs. speaker-oriented parentheticals

Modern scholars suggest various terms for the distinct types of parenthetical clauses. Kaltenböck (2007: 42) distinguishes two sub-types of parentheticals based on the verb's semantic category: *comment parenthetical clauses*, which make use of verbs of thinking in first- and second-person present tense. *Reporting parenthetical clauses*, which make use of message-conveying verbs, in third person and not limited to present tense.

Reinhart (1983: 175–176) introduced the terms *parenthetical-subject oriented* and *parenthetical-speaker oriented*. The first can be used as an answer to the question *What did the parenthetical-subject say or believe?* while the second can answer a much wider range of questions on the subject matter of the main clause.

Referring to examples (1)–(4), we see that the parenthetical clauses which have no syntactic relation (examples (1) and (3)) display the following features:

- The verbs are in third person.
- The clauses answer the question *What is said about the parenthetical subject?*
- The parenthetical clauses are objective and informative propositions.

Examples (2) and (4) on the other hand, have a syntactic relation and they

- include verbs and pronouns in second person;
- convey the speaker's attitude to the content of the utterance; and in some cases they express judgment regarding the host's subjects.

Having established the criteria for the distinction between two types of parenthetical clauses in the Qur'ān, we next discuss the factors of the positional flexibility of parentheticals.

3. Context and relevance as factors determining the position of parentheticals

Both types of parentheticals above are related to a particular element of the host. I argue that they must be located right after this element even if it interrupts the syntactic string of the host sentence. As I will show, a correct interpretation of the utterance depends on the precise place of the parenthetical clause. If it is placed elsewhere it will lose its relevance, namely, its contextual effect, and this might cause the listener to misinterpret the speaker's intended meaning. Since one of the main goals is to explain the position of parenthetical clauses, it requires the consideration of various language factors involved in planning and producing the information, such as the reception, decoding and the interpreatation of the message. The Relevance Theory provided by Sperber and Wilson (1996) explains how these factors work in communication, and, as will be shown, this theory helps to interpret the pragmatics behind the use of parenthetical clauses in the Qur'ān. Thus, I shall first provide some background information regarding this theory and then some core principles of this theory applied in my analysis will be explained.

Relevance Theory is a cognitive pragmatics theory of human communication. For Sperber and Wilson, people developed an ability to maximize the relevance of the utterance/gesture that they process. Since the listeners cannot pay attention to all information that reaches them, people had to develop some linguistic tools which enable them to focus the attention of the listener, to select it and to interpret the intention of the speaker.[21]

Most of the studies of grammar which take the Relevance Theory as the theoretical framework take a dynamic and inference-centerd approach in which grammatical structures are supported or refuted according to contextual constraints. Namely, a pragmatic and context-centered view of grammar is proposed in order to explain the choice of a specific grammatical structure and to explain how it contributes to a correct comprehension of the compound sentence.[22] Considering the parentheticals, instead of the typical approach, which tends to explain the type of information expressed by them (explanatory, emphatic etc.), parentheticals can be regarded as constraints on relevance – namely, they guide

[21] Yus (2009: 753–754).

[22] Yus (2009: 768).

the listener toward a correct comprehension of the compound sentence, since they reduce the effort needed to access the correct interpretation.[23]

In order to put forward a full framework of communication and show the nature of communication, Sperber and Wilson (1996) provided fundamental principles. In what follows, the principles televant to this study are presented:

(a) **Optimal relevance** is defined in terms of the cognitive effect and processing effort of the information. Thus optimal relevance is achieved when the effort invested in the information process decreases but the communicative benefit increases (Sperber and Wilson 1996: 48). The insertion of a parenthetical clause may appear to burden the interpretive process, but I suggest that it eases processing the information because it helps the audience to recover the intention behind the utterance without too much effort.[24]

(b) **Ostention** means making manifest an intention or showing someone something.

It occurs "when the communicator produces a stimulus which makes it mutually manifest to communicator and audience that the communicator intends, by means of this stimulus, to make manifest or more manifest to the audience a set of assumptions" (Sperber and Wilson 1996: 49, 63). In our case, the informative intention is linguistically communicated,[25] namely, the linguistic phenomenon of breaking the linear order of the sentence by inserting a parenthetical clause is a case of ostention. This structure is used to make the audience pay attention to this irregular structure and to understand the intention behind it.

(c) **Contextual effect** means modification and improvement of a context by using information that exercises some effect on that context. The two types of contextual effects are new information providing new evidence, thereby strengthening old assumptions, and information that provides evidence against old assumptions. An assumption which has no contextual effect in a given context is irrelevant (Sperber and Wilson 1996: 109, 121–125). The contextual effect significantly impacts the parenthetical's position. The speaker, when making an utterance and thinking of the contribution of the various constructions to the context, must be sure to keep the structures in their right places or else they will have no effect on the context. I now develop the effect of the Qur'ānic

[23] Yus (2009: 768) mentions this explanation when he refers to discourse markes and how they are viewed in Relevance Theory.

[24] Kavalova (2007: 167) suggests the same regarding *and-parentheticals*: "The insertion of an additional element in the string of the utterance may be seen as burdening the process of interpretation because the processing cost is increased (…) *and*-parenthetical clauses assist the achievement of certain cognitive effects which would be otherwise be less of fully inaccessible to the listener to process."

[25] For the relation between linguistic form and relevance, see Sperber and Wilson (2013: 152).

parenthetical clauses, and explain how they give rise to an efficient inferential process.

3.1. parenthetical clauses as strengtheners of God's omnipotence

(5)
fa-'in lam taf'alū wa-lan taf'alū fa-ttaqū n-nāra llatī waqūduhā n-nāsu wa-l-ḥiǧāratu 'u'iddat li-l-kāfirīna (Q 2:24)

"And if you do not – and you will not – then fear the fire, whose fuel is men and stones, prepared for unbelievers."

Q 2: 23 tells **the unbelievers that if they are in doubt about what God sent down to His servant Muḥammad, they are asked to adduce [at least] one Sura like those that God sent down** in the Qur'ān. The speaker (God), seeking to reject this ability by affirming that the unbelievers are unable to produce such a Sura, places the parenthetical in middle position. Hence the listener is aware that the speaker (God) not only refers to the possibility of producing Suras by the believers, but He (who knows all things – past, present and future) completely eliminates this possibility. By interrupting the main string with the parenthetical, the listener needs minimal effort to process this information and identify the intention behind the utterance. To prove this argument, we may read Q 2: 24, where the parenthetical clause is located at the end: **fa-'in lam taf'alū fa-ttaqū n-nāra llatī waqūduhā n-nāsu wa-l-ḥiǧāratu 'u'iddat li-l-kāfirīna wa-lan taf'alū* (Q 2: 24) "And if you do not then fear the fire, whose fuel is men and stones, prepared for unbelievers and you will not do so." The conditional sentence leads the listeners to understand that the possibility of creating a Sura exists, but if they do not do so they will suffer punishment. Yet when reading the last part of the sentence, the listener understands that his/her interpretation does not yield the speaker's intention (hence his/her processing efforts were wasted) and he/she should process the information again.

(6)
qul 'a-ra'aytum šurakā'akumu llaḏīna tad'ūna min dūni llāhi 'arūnī māḏā ḫalaqū mina l-'arḍi 'am lahum širkun fī s-samāwāti (Q 35: 40)[26]

"Have you seen your associates on whom you call, apart from God? Show me what they have created on the earth; or have they a partnership in the heavens?"
Regarding Q 35: 40, the parenthetical clause is formed as a command, which of course cannot be fulfilled by the polytheists. This is done to prove to the polytheists that their faith is wrong by emphasizing that their idols are powerless and that the Creator of heaven and earth is the only God, and there will be no

[26] Similar cases are Q 2: 80; 41: 15; 30: 4

object of worship except Him. Had the parenthetical clause been introduced at the end, the listener could not have drawn any conclusion regarding the reason why the idols, whom they worship, cannot be associated with or likened to God. The absence of the parenthetical (or its placement elsewhere) triggers the assumption that the verse expresses a regular question which does not carry any admonition concerning the idols and the polytheists.

(7)
wa-mā n-naṣru 'illā min 'indi llāhi l-'azīzi l-ḥakīmi li-yaqṭa'a ṭarafan mina llaḏīna kafarū 'aw yakbitahum fa-yanqalibū ḫā'ibīna laysa laka mina l-'amri šay'un 'aw yatūba 'alayhim 'aw yu'aḏḏibahum fa-'innahum ẓālimūna (Q 3: 126-128)[27]

"Help comes only from God the Almighty, the all-wise; He will cut off a part of the unbelievers or frustrate them, so that they turned in their tracks, disappointed, no part of the matter is yours [you can do nothing against this], or He will turn toward them again, or chastise them, for they surely are evildoers."

There are two ways to interpret the parenthetical clause in example (7):[28] first, this clause starts a new sentence hence the utterance may be interpreted as: "It is not for you to decide whether He will accept their repentance, or chastise them for they surely are wrongdoers." Second, as a parenthetical clause placed in middle position between two connected clauses which can be interpreted as: "God provided His aid to you in order to cut off a part of those who disbelieved and frustrate them so that they retreat in disappointment –no part of the matter is yours – or He will accept their repentance, or chastise them, for they surely are wrongdoers." The intention behind verses 126–128 is to clarify that a triumph (in the battlefield), defeat of the unbelievers, forgiveness and punishment—all depend on God's will. Not only do people take no part in such actions, they can not do anything against God's decision. This idea is established in verse 126 (*wa-mā n-naṣru 'illā min 'indi llāhi*) and the parenthetical clause clearly manifests this intention. Still, a reservation must be made against this explanation because in this case, the same kind of contextual effect could be possible were the parenthetical clause placed at the end. The listener is informed that God cut off a part of the unbelievers or frustrated them; he turns toward them, or punishes them, and at the end the parenthetical clause states that whatever the case, the people can do nothing about it.

In the next example the parenthetical clause is inserted where it is most relevant because it strengthens a previous statement declaring that the people beg forgiveness from God when they commit an indecency. They do so because the only one who can forgive their deeds is God:

[27] A similar case is Q 30:2

[28] See, e.g., Rāzī (2000: vol. 7–8: 191).

(8)
wa-llaḏīna ʾiḏā faʿalū fāḥišatan ʾaw ẓalamū ʾanfusahum ḏakarū llāha fa-staġfarū li-ḏunūbihim ***wa-man yaġfiru ḏ-ḏunūba ʾillā llāhu*** *wa-lam yuṣirrū ʿalā mā faʿalū wa-hum yaʿlamūna* (Q 3: 135)
"Those who, when they commit an indecency or do injustice to themselves, remember God, and seek for forgiveness for their sins – **and who forgives sins except God**? – and do not knowingly persist in what they have committed."

3.2. Elucidating parentheticals

(9)
wa-llaḏīna ʾāmanū wa-ʿamilū ṣ-ṣāliḥāti ***lā nukallifu nafsan ʾillā wusʿahā*** *ʾulāʾika ʾaṣḥābu l-ǧannati hum fīhā ḫālidūna* (Q 7: 42)[29]
"And those who believe, and do deeds of righteousness – **we do not impose upon a soul [duties] but only according to its capacity** – those are the inhabitants of Paradise, therein dwelling forever."

The idea that Paradise is the reward of faith and good deeds occurs repeatedly in the Qurʾān. For example, Q 2: 82 has the same structure as Q 7: 42 except for the parenthetical. If the listener is already exposed to this idea it might be argued that s/he can easily process the information in Q 7: 42. However, this old information is connected to new information, expressed in the parenthetical clause, and this connection produces more new information, which can be derived by inference.[30] Thus, the listener may infer that the speaker's aim is not simply to communicate that those who believe and do good deeds will be rewarded, but also to restrict the definition of what good deeds are. The fulfillment of all good deeds is the best, but God knows that it might be a burden upon the people. Therefore, throughout the parenthetical clause He clarifies that every one is responsible for as many of the good deeds s/he is able to perform, and will still be rewarded.

(10)
wa-la-ʾin ʾaṣābakum faḍlun mina llāhi la-yaqūlanna ***ka-ʾan lam takun baynakum wa-baynahu mawaddatun*** *yā-laytanī kuntu maʿahum fa-ʾafūza fawzan ʿaẓīman* (Q 4: 73)
"And if some grace from Allah befall you, He will surely say – **as if there had not been any affection between you and Him** – I wish I had been with them to attain a mighty triumph!"

Verses 72–73 describe the behavior of the hypocrites. When the believers return from the battlefield having suffered a failure or martyrdom, the hypocrites will happily say: What great bounty God has given them that they did not accompany the believers to witness their defeat and to suffer as they suffered. Yet as soon as the hypocrites are informed that the believers have gained the victory, and naturally have won some booty, they feign contrition, saying if only they had been with the believers to achieve a great triumph. The question is what is the

[29] A similar example is Q 18: 30-31.
[30] For old and new information in a process of inference, see Sperber and Wilson (1996: 48).

contextual effect of the parenthetical clause in this context? According to Rāzī (1993: vol. 5, 186), the parenthetical clause expresses wonder (*ta'aǧǧub*) about the (absurd) behavior of the hypocrites; even though there is no affection between the hypocrites and the believers, the hypocrites are willing to associate with the believers when it comes to booty.

(11)
wa-waṣṣaynā l-'insāna bi-wālidayhi ***ḥamalathu 'ummuhu wahnan 'alā wahnin wa-fiṣāluhū fī 'āmayni*** *'ani škur lī wa-li-wālidayka 'ilayya l-maṣīru* (Q 31: 14)
"And we did enjoin upon man concerning his parents – his mother bore him in weakness upon weakness, and his weaning takes two years, (saying:) 'Be thankful to Me and to your parents: unto Me is the ultimate return."

In Q 31: 14 people are commanded to show kindness and gratitude to their parents. Note, however, that the parenthetical clause refers only to the mother, while in the host sentence both parents are mentioned. The mother feels weak and the weakness gets worse in the course of pregnancy[31] because of natural changes that affect her body. Her tired physical condition lasts almost three years (the course of pregnancy and the suckling). According to the literature, the parenthetical clause emphasizes the mother's travails so as to attract peoples' attention to the mother's physical and emotional sacrifice during the pregnancy and suckling (Ḥasan 2012: 239). The problem is that such an argument refers to the parenthetical's content, and it fails to explain how it is linked to the host or how it contributes to understanding the complete verse. I would suggest two possible explanations for the insertion: first, the parenthetical is inserted after the opening clause to make the command relevant to all people. From reading the first part *wa-waṣṣaynā l-'insāna bi-wālidayhi*, it might be possible to assume that some people, who have cut their relations with their parents for familial or personal issues, or are constantly at loggerheads with their parents, might think that this command does not concern them. However, the parenthetical clause reminds those people that whatever the circumstances, they owe their lives to their mothers. For this reason they should be thankful to her and to the father, who accompanied the mother during her pregnancy. Second, in this case the parenthetical clause and the host are both required as input into the inferential process, which can be schematized as follows:

- Presentation of the command: The people are commanded concerning their parents.
- Why it is commanded: because the mother suffered during the pregnancy.
- What they are commanded: to be thankful to both the mother and the father.

Thus, the inferential process yields that the intended cognitive effect of this utterance is to command, but also to explain the reason for giving this command.

[31] For this explanation, see Zamaḫšarī (1947: 3, 494–495).

Finally, I go back to example (2)[32] (*'inna llaḏīna kafarū* ***sawā'un 'alayhim 'a-'anḏartahum 'am lam tunḏirhum*** *lā yu'minūna* Q 2: 6). It should be mentioned that some commentators, such as Rāzī (2000: vol. 1–2, 38) point to two options for analyzing Q 2: 6:

1) *llaḏīna kafarū* =subject; *lā yu'minūna* = predicate. According to this analysis, *sawā'un 'alayhim 'a-'anḏartahum 'am lam tunḏirhum* functions as a parenthetical clause and interpolates between the subject and the predicate.
2) *llaḏīna kafarū* =subject (*'ism 'inna*); *sawā'un* = predicate (*ḫabar 'inna*); *lā yu'minūna* = new sentence. In this option there is no parenthetical clause.

The insertion of the parenthetical clause in Q 2: 6 is crucial for the inferential process.

Placing the parenthetical at the end leaves little information to process. Stating that those who are infidels do not believe in God is an obvious declaration that requires no information processing. However, when the parenthetical clause is inserted into the host sentence the utterance starts from a set of premises and ends with the conclusion that unbelievers, warned or not, will always have the same (false) beliefs and the same (bad) habits.

(12)
fa-lammā waḍa'athā qālat rabbi 'innī waḍa'tuhā 'unṯā wa-llāhu 'a'lamu bi-mā waḍa'at ***wa-laysa ḏ-ḏakaru ka-l-'unṯā*** (3: 36)[33]
"And when she gave birth to her, she (the wife of Imran) said: Lord, I have given birth to her, a female, while God knew very well what she had given birth to; **the male is not as the female**."

In Q 3: 36 the commentators identify the clause *wa-llāhu 'a'lamu bi-mā waḍa'at* as parenthenthetical; [34] however, I argue that this clause functions as a circumstantial clause while the parenthetical clause is *wa-laysa ḏ-ḏakaru ka-l-'unṯā* for the following reason: in verse 35 we read that Imran's wife says to God that she **vowed to Him to dedicate her child to Him.** In verse 36 she says to God that she gave birth to a girl, but God already knew that. The question that arises here is why does she state that she has a daughter if God already knows that? The answer is found in the parenthetical clause, which explains that a male is not like a female, hence the audience may infer that the clause *'innī waḍa'tuhā 'unṯā* expresses regret and disappointment, because Imran's wife had hoped to bear a son. A male is not like a female because he can devote himself completely to the

[32] Examples one and three can also be classified under this category; their contextual effect is clarified in section 2.
[33] A similar case is Q 4:25.
[34] See e.g., Ibn 'Āšūr (1992: vol. 3–5, 233); Ṭanṭāwī (1992: vol. 2, 87).

service of God in the place of worship, and a female cannot fulfill the rituals as a male can.[35]

3.3. Parenthetical clause as a qualifier

(13)
wa-llaḏīna 'āmanū wa-'amilū ṣ-ṣāliḥāti wa-'āmanū bi-mā nuzzila 'alā muḥammadin ***wa-huwa l-ḥaqqu min rabbihim*** *kaffara 'anhum sayyi'ātihim wa-'aṣlaḥa bālahum* (Q 47: 2)
"But those who believe and do righteous deeds and believe in what is sent down to Muḥammad – **and it is the truth from their Lord** – He will absolve them of their evil deeds and dispose their minds aright."

Q 47:2 also speaks of the reward for those who believe and do good deeds, but it adds that one has to believe in the Qur'ān which was sent down to Muḥammad. Now the question is why is the parenthetical clause inserted between the subject and the predicate? It is introduced to modify the Qur'ān by saying that it is true, unchangeable, and steadfast.[36] The qualifier is placed immediately after the qualified component (*bi-mā nuzzila 'alā muḥammadin*). In this case, disconnection of the qualifier from the qualified noun will generate an ungrammatical structure, but also an incomprehensible utterance.

(14)
wa-'ini mra'atun ḫāfat min ba'lihā nušūzan 'aw 'i'rāḍan fa-lā ǧunāḥa 'alayhimā 'an yuṣliḥā baynahumā ṣulḥan ***wa-ṣ-ṣulḥu ḫayrun*** (Q 4: 128)
"And if a woman fears ill treatment from her husband, or desertion, there is no sin on them if the couple set things right between them, and right settlement/reconciliation is better."

Q 4: 128 states that it is better for the spouses to come to a mutual understanding so that the wife may remain with her husband. The parenthetical clause qualifies this situation as the best for both sides.

(15)
ittabi' mā 'ūḥiya 'ilayka min rabbika ***lā 'ilāha 'illā huwa*** *wa-'a'riḍ 'ani l-mušrikīna* (Q 6: 106)
"Follow what has been revealed to you from your Lord, **there is no God but He**, and turn away from idolaters."

---

35 For this explanation of *wa-laysa ḏ-ḏakaru ka-l-'unṯā*, see Rāzī (2000: vol. 7–8, 24).

36 Bayḍāwī (1996: vol. 1, 190), who identifies *wa-huwa l-ḥaqqu min rabbihim* as a parenthetical clause, explains that the Qur'ān is *ḥaqq* for being *nāsiḫ lā yunsaḫ* "[a book that] abrogates and is not abrogated."

The parenthetical clause in Q 6: 106 qualifies the noun *rabb* "Lord," but at the same time it explains the logic/reason behind the two commands, as inferred from Rāzī's explanation:

> ʻalā ʼannahu taʻālā lammā kāna wāḥidan fī l-ʼilahiyyati fa-ʼinnahu yağibu ṭāʻatuhu wa-lā yağūzu l-ʼiʻrāḍi ʻan takālīfihi bi-sababi ğahli l-ğāhilīna wa-zayġi z-zāʼiġīna[37]
> "Since God sublime is the only [real God among the divine entities] he [the prophet Muḥammad] must be obeyed and avoidance of fulfilling His commandments is impermissible because of the ignorance of the ignorant and the deviation of the deviators."

3.4. Parenthetical clauses as background information

In this section the parentheticals are relevant in the sense that they provide important information for understanding later developments in the story. This account is a departure from that proposed by Zewi (2007: 67), who argues that in Biblical Hebrew parentheticals introduce~~s~~ background information. She defines this term thus:

> "[Background information in this book] restricted to extra information inserted into the story by the scribe or narrator to facilitate understanding of certain developments that could not be understood without it."

(16)
wa-ğāʼahū qawmuhu yuhraʻūna ʼilayhi ***wa-min qablu kānū yaʻmalūna s-sayyiʼāti*** *qāla yā-qawmi hāʼulāʼi banātī hunna ʼaṭharu lakum ʼa-laysa minkum rağulun rašīdun* (Q 11: 78)
"And his people came rushing toward him; and previously they had been doing evil deeds. He said: 'O my people, here are my daughters – they are purer for you. So fear Allah, and do no degrade me before my guests! Is there not among you one upright man?"

Verse 77 states that when God's messengers came to Lot, he grieved for them and felt constrained to protect them. In verse 78 we discover that he worries about his people because they have committed evil deeds in the past. Furthermore, the parenthetical clause helps in understanding why Lot offers his daughter for marriage and then he asks them not to degrade him before his guests. Having known at the early stage of the utterance that the people are used to committing sins, the listener infers that Lot, throughout his words and actions, wants to prevent his people from doing what is forbidden.

[37] Rāzī (2000: vol. 13–14, 113).

(17)
'inna qārūna kāna min qawmi mūsā ***fa-baġā 'alayhim*** *wa-'ātaynāhu mina l-kunūzi mā 'inna mafātiḥahu la-tanū'u bi-l-'uṣbati 'ulī l-quwwati 'iḏ qāla lahu qawmuhu lā tafraḥ 'inna llāha lā yuḥibbu l-fariḥīna* (Q 28: 76)
"Verily Korah was of the people of Moses, and he oppressed them; and We had given him of the treasures so much that its keys would have been a burden to a company of men endowed with strength. When his people said to him: Do not exult, God does not love those that exult."

In Q 28: 76 we are introduced to Korah and we are informed that he belongs to the people of Moses. Then the parenthetical clause provides further information: the people have suffered injustice at his hand. In the last part of the verse he is described as a rich, arrogant man – a wealthy man, indeed the manifestation of wealth, accompanied by pride and deceit. Reading the parenthetical clause the listener cannot but wonder how a man who is negatively presented becomes so rich, and instead of being punished for his pride and his ignorance of the people's social condition, their need of help, he is still being rewarded. This contradiction is required as input into the inferential process that yields the real intention (or message) of the story of Korah. It shows the people that they can have wealth and property; the important thing is the way it is used. If it is to show pride, negligence, injustice and oppression, then the man will be punished. Furthermore, **people are usually dazzled by the wealth and the pleasures offered in this world, while modesty,** piety, and righteous deeds are the most important things that guarantee well-being in the hereafter.

4. Conclusion

Parenthetical clauses can be placed at the middle or final position. The middle-positioned parentheticals are syntactically integrated into the host by an anaphoric relation, while the final-positioned parentheticals are syntactically unrelated to the host.

It is claimed that parenthetical clauses interrupt the prosodic flow of the sentence by introducing intonational breaks in the host sentence; however, their insertion is a communicative act for reaching an optimal relevance. Namely, the addressees do not waste much efforts in the interpretation of the Qur'ānic verse. How is the principle of optimal relevance satisfied throught the insertion of the parenthetical clause?

The parenthetical's position is determined by the place of the component to which it refers. Only when it takes this position it has a contextual effect, which can be immediately identified by the addressees and hence he interprets the utterance correctly. Qur'ānic arenthetical clauses has various contextual effects: They serve to affirm God's omnipotence, indicating that only God produces Suras, created heaven and earth. He is the forgiver and all depends on His will; to explain a statement, for example, when people are required to do righteousness, the

parenthetical clause explains that it must be done in accordans to one's capability to perform righteousness; to qualify, to highlight a specific characterization, for example, one of the parenthetical clauses modifies the Qur'ān as the truth from God; to provide background information, which could explain further developments in the narrative.

If the parenthetical clauses would have been located elsewhere in the utterance then they would have been irrelevant, without a clear effect.

Scholars argue that the host is in no way grammatically dependent on the parenthetical clause and therefore it can be omitted. However, establishing the contextual effects of the Qur'ānic parenthetical clause can provide convincing evidence that parenthetical clauses do not carry peripheral information but contribute to a correct interpretation of the host; therefore, they cannot be dropped. The representation of the intended utterance is achieved by the fact that despite the interpolation the information is arranged in logical and sequential order, providing the listener the prerequisites that guarantee an efficient inferential process. In an inferential process, the listener should proceed from one premise to another, ending with a conclusion. This process can be demonstrated by example (1), which consists of three premises: first, a condition (if you do); second, the speaker's comment (you will not do); third, the result (you will be punished). The conclusion drawn by the listener is that the polytheists are definitely not able to produce a Sura, therefore they will be punished. Changing this sequenced order can create an utterance where the listener derives a certain intention even before reaching the end of the utterance; but then he/she might encounter a parenthetical clause that will require re-processing all the information in the utterance.

References

Arberry, Arthur J. 1964. *The Koran Interpreted*. London: Oxford University Press.

Burton-Roberts, Noel. 2006. Parentheticals. In: Keith E. Brown (ed.), *Encyclopaedia of Language and Linguistics*, 179–189. Amsterdam: Elsevier.

Dehé, Nicole, Kavalova Yordanke. 2007. Parentheticals: An Introduction. In: Nicole Dehé and Yordanka Kavalova (eds.), *Parentheticals*, 1–25. Amsterdam/Philadelphia: John Benjamins.

El-Awa, Salw. 2006. *Textual Relations in the Qur'ān: Relevance, Coherence and Structure*. London and New York: Routledge.

Gully, Adrian. 1995. *Grammar and Semantics in Medieval Arabic: A Study of Ibn-Hisham's Mughni l-Labib*. London: Curzon Press.

Haegeman, Liliane. 1988. Parenthetical Adverbials: The Radical Orphanage Approach. In Satoshi. Chiba (ed.), *Aspects of Modern English Linguistics*. Tokyo: Kaitakushi, 232–254.

Hoffmann, Ludger. 1998. *Parenthesen*. *Linguistische Berichte* 175. 299–328.

Jackendoff, Ray. 1977. *X' Syntax: A Study of Phrase Structure*. Cambridge, MA: The MIT Press.

Kaltenböck, Gunther. 2007. Spoken Parentheticals Clauses in English: A Taxonomy. In: Nicole Dehé and Yordanka Kavalova (eds.), *Parentheticals*, 25–53. Amsterdam/Philadelphia: John Benjamins.

Kavalova, Yordanke. 2007. And-Parenthetical Clauses. In: Nicole Dehé and Yordanka Kavalova (Eds.), *Parentheticals*, 145–173. Amsterdam/Philadelphia: John Benjamins.

Peterson, Peter. 1999. On the Boundaries of Syntax. In: Peter Collins and David Lee (eds.), *The Clause in English*, 229–250. Amsterdam: John Benjamins.

Potts, Christopher. 2002. The Syntax and Semantics of As-Parentheticals. *Natural Language and Linguistic Theory*, 20(3). 623–698.

Reckendorf, Hermann. 1921. *Arabische Syntax*. Heidelberg: Carl Winter's Universitätsbuchhandlung.

Reinhart, Tania. 1983. Point of View in Language: The Use of Parentheticals. In: Gisa Rauh (ed.), *Essays on Deixis*, 169–194. Tübingen: Narr.

Sperber, Dan, Wilson, Deirdre. 1996. *Relevance: Communication and Cognition*, 2nd ed. Oxford: Blackwell.

Sperber, Dan, Wilson, Deirdre. 2013. Linguistic Form and Relevance. In: Dan Sperber, Deidre Wilson (Eds.), *Meaning and Relevance*, 149–169. Cambridge: Cambridge University Press.

Yus, Francisco. 2009. Relevance Theory. In: Heine Bernd and Narrog Heiko (eds.), The Oxford Handbook of Linguistic Analysis, 753–777. Oxford and New York: Oxford University Press

Zewi, Tamar. 2007. *Parenthesis in Biblical Hebrew*. Leiden and Boston: Brill.

Ziv, Yael. 1985. Parentheticals and Functional Grammar. In: Machtled A. Bolkestein et al. (eds.), *Syntax and Pragmatics in Functional Grammar*, 181–213. Dordrecht: Foris Publications.

Sources in Arabic

al-'Askarī, 'Abū Hilāl al-Ḥasan Ibn 'Abd Allāh. 1971. *Kitāb aṣ-ṣinā'atayni: al-kitab wa-š-ši'r*. Beirut: Dār al-kutub al-'ilmiyya.

al-Bayḍāwī, 'Abdallāh Ibn 'Umar. 1996. *'Anwār at-tanzīl wa-'asrār at-ta'wīl*. Cairo: Maktabat wa-maṭba'at al-mašhad al-ḥusaynī.

Ibn Ǧinnī, 'Abū al-Fatḥ 'Uṯmān. 1913. *al-Ḫaṣā'iṣ*. Cairo: Maṭba'at al-hilāl.

Ḥasan, Sāmī 'Aṭā. 2012. *al-Ǧumla al-mu'tariḍa fī l-Qur'ān al-karīm: mawāḍi'uhā wa-dalālatihā*. Amman: Dār al-furqān li-n-Našr wa-t-tawzī'.

Ibn 'Āšūr, Muḥammad aṭ-Ṭāhir. *at-Taḥrīr wa-t-tanwīr*. 1992. ad-Dār at-tunisiyya li-n-našr.

'Ibn Hišām, Ǧamāl ad-Dīn al-'Anṣārī. 1991 *Muġnī al-labīb*. Beirut: Dār al-Ǧīl.

an-Nasafī, 'Abdallāh Ibn Aḥmad. 1996. *Madārik at-tanzīl wa-ḥaqā'iq at-ta'wīl*. Beirut: Dār an-nafā'is.

al-Qazwīnī, Ǧalāl ad-Dīn Muḥammad Ibn 'Abd ar-Raḥman. 2002. *Talḫīṣ al-miftāḥ fī l-ma'ānī wa-l-bayān wa-l-badī'*. Beirut: al-Maktaba al-'aṣriyya.

ar-Rāzī, Muḥammad Faḫr ad-Dīn. 1993. *Mafātīḥ al-ġayb*. Beirut: Dār al-fikr li-ṭ-ṭibā'a wa-n-našr wa-t-tawzī'.

ar-Rāzī, Muḥammad Faḫr ad-Dīn. 2000. *Mafātīḥ al-ġayb*. Beirut: Dār al-kutub al-'ilmiyya.

aš-Šawkānī, Muḥammad 'Alī Ibn Muḥammad. 1997. *Fatḥ al-qadīr*. al-Manṣūra: Dār al-wafā' li-ṭ-ṭibā'a wa-n-našr wa-t-tawzī'.

Ṭanṭāwī, Muḥammad Sayyid. 1992. *at-Tafsīr al-wasīṭ li-l-Qur'ān al-karīm*. Cairo: Dār al-ma'ārif.

Yāqūt, Maḥmūd Sulaymān. 1998. *'I'rāb al-Qur'ān al-karīm*. Dār al-ma'rifa al-ǧami'iyya.

az-Zamaḫšarī, 'Abū al-Qāsim Muḥammad Ibn 'Umar. 1947. *al-Kaššāf 'an ḥaqā'iq ġawāmiḍ at-tanzīl*. Beirut: Dār al-kitāb al-'arabī.

BACK TO ORTHOEPIA – SPELLING IN PRONUNCIATION INSTRUCTION: "WORDS COMMONLY MISPRONOUNCED" BY LEARNERS OF SIX L1S*

MARTA A. NOWACKA
University of Rzeszow
martha.nowacka@gmail.com

Abstract
This is a continuation of Nowacka's (2016) study on the importance of local and global errors and spelling in pronunciation instruction. Unlike in the previous research that focused on the performance of Polish learners only, respondents of six different nationalities are included, in search of some cross-national universals or absence of them.

This study seeks to answer the following questions: whether there is a need to focus on spelling in a pronunciation course with learners representing six different L1s and if this is the case which graphophonemic / phonotactic rules of English should be explicitly taught to all of these learners and which ones might be L1 specific only.

The intention is also to empirically confirm the existence of local errors in the performance of around 240 speakers and 50 more listeners, constituting 291 listeners of six nationalities (Kazakh, Malaysian, Polish, Turkish, Tajik and Ukrainian) and to confirm the usefulness of memorizing Sobkowiak's (1996) 'Words Commonly Mispronounced' even for learners of different L1s.

Keywords: graphophonemic rules, letter-to-sound-correspondence, pronunciation instruction, spelling, 'Words Commonly Mispronounced'

1. Introduction

1.1. The correspondence between spelling and pronunciation

> The issues of orthoepia and to some extent the problems of orthography are connected with the study of phonetics … in the English language there are great discrepancies between a written version and its sound form … to some degree English spelling is of ideographic kind. However, English spelling is not completely ideographic. It is only, especially from a descriptive perspective, very inconsistent, and therefore not very phonetic.[1] (Jassem 1971: 65)

* I would like to express my deepest thanks to around 300 anonymous students of thirteen nationalities from the University of Rzeszów and University of Information Technology and Management in Rzeszów, Poland who voluntarily completed a questionnaire and participated in a recording session, without whom this paper would not have come into existence.

1 Translation of this quote from the Polish language is mine.

In this section we intend to provide arguments for the inclusion of spelling-oriented activities, with a focus on spelling-to-sound relations and also spelling irregularities in pronunciation training as standard techniques to improve foreign learners' competence and then performance which should accompany the formation of English sound inventory and suprasegmentals.

Since the focus of this paper is spelling and its relation to pronunciation we find it appropriate to present the reader with a short background of the English spelling system. According to the English Spelling Society's report, English has 185 graphemes[2] for 44 sounds, as opposed to the European average of 50 spellings, which means that it is not a completely regular system and therefore learning to read and write it is more difficult than other alphabetic writing systems in which there is nearly a one-to-one relationship between their sounds and spellings, usually with spellings outnumbering sounds by just a few.

> The basic English spelling system has 91 patterns … 80 main spellings, 8 for unstressed endings, 2 prefixes and the consonants doubling rule. … 80 English spelling patterns are undermined by one or more alternatives, e.g. [c]at – pl*ai*t, mer*i*ngue …. The greatest English reading difficulties, however, are caused by the 69 spellings which have more than one pronunciation … They make at least 2000 English words not completely decodable. (English Spelling Society 2017a)

Carney (1994: 18) observes that in the English writing system instead of mapping phonemes on to letters we usually keep the spelling of a morpheme constant in spite of the varying pronunciation of the morpheme in different contexts, e.g. in *mime* vs. *mimic, sane* vs. *sanity, cone* vs. *conical*, and children learn both a long and a short phonemic value for the simple vowel letters, i.e. the letter <i> can stand for /aɪ/ or /ɪ/, <a> = /eɪ/ or /æ/, <o> = /əʊ/ or /ɒ/ as in the above-mentioned examples. Interestingly only in the pair of phonemes /aʊ/ - /ʌ/ do the vowel letters in spelling sometimes vary to reflect the surface difference, e.g. in *pronounce* vs. *pronunciation* but not in *south* vs. *southern*.

The 109-year-old English Spelling Society (2017b), formerly the Simplified Spelling Society, aims at promoting spelling reform and raising awareness of the problems and costs caused by the irregularity and complexity of English spelling. It has formulated six axioms on English spelling from which we learn that the alphabetic principle in the English spelling system is breeched, i.e. the letters of the alphabet, which were originally designed to represent speech sounds, do not perform their primary function well due to on-going changes in pronunciation but not in spelling over the period of the last 1,000 years. Letters do not make it easy for the reader to pronounce words from their written form, and for the writer to spell words when heard, which calls for simplification to reduce the strain on the part of language users, who need more time to master this system, and for the society as a whole because of the longitudinal costs.

[2] Bell (2009) points to 205 graphemes in the English spelling system.

Students have to reconcile what they say in English with how it is written. The Spelling Society website admits that English children in a naturalistic setting also struggle with associating letters with a particular sound and it takes them up to three years longer to eventually master it than it takes for those who use other alphabetic languages.

Foreign learners of English learn new words not only from the oral but also to a great extent from written sources. This leads us to try to guess the pronunciation of a word from the orthography, which is frequently not an easy task even for native-speakers. Carney (1994: 31) comments on this issue in the following way:

> literacy for foreign learners of English is a special case. In the absence of a live informant, the orthography is the main indication available of the pronunciation of an unfamiliar word, short of looking it up in a pronouncing dictionary. They need, from the very beginning, some awareness of spelling-phoneme correspondences as a key to the phonetics of English and to prevent them from being misled by the writing system of their first language, in which they will usually be literate.

Jones (1932: 7) observes that "[t]he result of such inconsistencies [between pronunciation and ordinary spelling] is that the foreigner who depends solely on ordinary orthography is in innumerable cases at a loss to know what sounds should be used, and is continually mispronouncing words."

Wells (2010, 2011c) explains the phenomenon of spelling pronunciation which happens when "[a] speaker who is familiar with the written form of a word but not with its spoken form may, on the basis of the spelling, infer a pronunciation different from the traditional or generally used one," e.g. in *backwards* as /ˈbækwədz/ which replaced the former /bækədz/ or *honorarium* as /ˌhɒnəˈreəriəm/.

Wells (2010) notes that there are two other phenomena related to the issue of spelling pronunciation. One of them is pronunciation spelling, sometimes called phonetic spelling, in which a new spelling represents the pronunciation better than the traditional one, for example that applied by Lewis (2017) in his phonetic blog, e.g. writing the word *said* as *sed* or *unstressed* as *unstrest* in the excerpt that follows "nobody: this word can be sed to have two strongforms … /ˋnoʊbɒdi / .. [i]s not used …when unstrest …" Lewis (2017). Wells (2010) also defines 'non-spelling pronunciation,' as "the adoption of a new pronunciation that does not match the traditional spelling," e.g. represented by the pronunciation variant /mɪsˈtʃiːviəs/ of *mischievous*, which might in the future lead to a change in spelling, e.g. *mischievious*.

We do not wish to spread the heresy that by familiarizing oneself with English spelling conventions one will assimilate all the necessary phonetic details. What we are arguing, after Carney (1994) and Wells (2011), is that being aware of spelling-to-sound correspondences is a necessary prerequisite for literate language users to read English correctly or to make informed guesses about how unfamiliar lexical items are pronounced. It is undeniable that good English pronunciation

involves acceptable performance, i.e. good articulation of English sounds and suprasegmentals combined with sufficient competence, including the skill of turning a variety of letter combinations from a written text into appropriate sounds. It appears that memorizing the pronunciation of some lexical items, whose orthography is far from transparent and whose spoken form is surprising in view of the spelling, is a routine procedure used by native and non-native learners (Bell, 2009, 2010a-b, 2015), e.g. pronouncing the diagraph <ei> irregularly as /iː/ in c*ei*ling, /aɪ/ in h*ei*ght and as /e/ in h*ei*fer, in other words, not in accordance with the suggestion from the spelling, i.e. rendering it as /eɪ/ as in *veil*.

Wells (2011b) reminds us that pronunciation training, apart from letter-to-sound rules, should also encompass the above-mentioned lexical spelling

> … teachers of English pronunciation need to give a lot of attention to establishing the correct target for the pronunciation of each word in the student's English vocabulary. Knowing spelling is not enough. We're all aware that the relationship between spelling and pronunciation is less than perfect. But we often don't realize how insidious the misleading effect of the orthography can be. Wild guesses are not the route to follow.

Wells (2012) gives the example of a fluent French speaker's mispronunciation of the word *idea* as ID /aɪˈdiː/ leading to some confusion, and sums up that

> [i]t shows that good pronunciation in EFL depends not just on being able to make the right sounds and phonemic contrasts, and to master syllable structure (clusters, final consonants etc.), but also to know the right pronunciation for every word in your vocabulary. … And not to be misled by the spelling. (Given *sea* **si**: and *flea* **fli**:, you can see the problem.)

Wells (2011a) is a keen supporter of spelling reform, welcoming the phonetically-based *bilding a cubbard* without automatically rejecting the traditional *building a cupboard*: "I don't see the logic in insisting that the traditional spellings must no longer be permitted alongside the reformed spellings. Why not allow the two forms to co-exist, to compete if you will, until one or other becomes obsolescent and ultimately obsolete?"

In yet another blog entry, Wells (2011b), in a comment on a phonemic transcription of units of measurements in an Italian edition of *The Pond Travel Kit Inglese*, which is full of mistakes, *e.g. 1 pint* as /wʌn pɪnt*/ instead of /wʌn paɪnt/ or 1 *ounce* as /wʌn oːnts*[3]/ instead of /wʌn aʊnts/, once again emphasizes the necessity of memorizing the pronunciation of individual words: "[i]t's not just a matter of learning to make the sounds of English in an acceptable way. It's also a matter of knowing which sounds ought to be used in which words. And that's what often gets neglected."

It is frequently stated in the literature that for the development of reading skills in English a growing awareness of sound and letter correspondences plays a significant role. Ellis and Cataldo's (1990 in Carney, 1994) results prove that

[3] Asterix stands for the erroneous pronunciation.

spelling is an important contributor to early reading; however, this interaction is unidirectional, which means that good reading does not predict good spelling.

The view we are taking in this study supports that one of Carney (1994: 32) who believes that "the identification of a word in reading is an informed guess and that several channels more-or-less simultaneously bring relevant information to bear, one of which channels may be spelling-to-sound correspondences." We believe that by presenting a systematic description of some regularities of English spelling, even those that are far from straightforward, and insisting on memorizing some lexical items with inconsistent pronunciation, we would increase our learners' competence and prepare them for making informed guesses about the rendition of unfamiliar words. Only then we might expect that regularity is in the eye of the beholder.

It is hoped that the rules we have selected for thorough practice in the classroom on the basis of the experiment might contribute to reducing the number of irregularities to be learnt on the part of a learner and in general it might lead to their acceptable, clearly understood pronunciation in English.

1.2. Classification of pronunciation errors

In the previous paper on the influence of English spelling on Polish learners' pronunciation (Nowacka, 2016) we discussed arguments for regularity of English spelling (Upward and Davidson, 2011; Crystal, 2012), spelling-induced mispronunciations in the latest Polish research (Sobkowiak, 1996; Scheuer, 1998; Majer, 2002; Szyszka, 2003; Szpyra-Kozłowska 2005, 2013, 2015; Szpyra-Kozłowska and Stasiak, 2010; Nowacka et al., 2011; Pęzik and Zając, 2012; Bryła-Cruz, 2013; Porzuczek, 2015; Waniek-Klimczak, 2015 and Zając, 2015) and also the notion of local and global pronunciation errors (Sobkowiak, 1996; Porzuczek, 2015; Szpyra-Kozłowska, 2015).

In Nowacka's (2016) study we implemented the classification of phonetic errors into local and global after Porzuczek (2015), which agrees with the taxonomy by Sobkowiak (1996) and Szpyra-Kozłowska (2015). To familiarize the reader with the terminology applied here we present Porzuczek's (2015: 172) definitions of global and local errors. Global errors, caused either by L1 interference or the learners' failure to follow the most characteristic pronunciation patterns, are reported to be easily avoided if typical spelling cues are taken into account. Unlike global mispronunciations, local ones, which are regarded as graphophonemic exceptions, can hardly be prevented by observing the graphophonemic or phonotactic rules of English. On the basis of this dual categorization Porzuczek (2015) distinguishes three major classes of erroneous pronunciation: the unavoidable local errors, further subdivided into true and 'either-or,' and avoidable globalized errors. Each group of these phonetic mispronunciations is further separated into 27 patterns, each referring to one aspect of English phonotactics and/or spelling-phonology relations.

It should be added that in the second- and foreign-language acquisition studies there are different classifications of language errors. For example, James (1998) introduces interlingual and intralingual errors. The former, mother-tongue induced, can be predicted and their causes can be determined, the latter, which according to Richards (1974) constitute 75% of all kinds of errors, are the result of misinterpretation of target language rules. They occur when learners engage their learning strategies and do one of the following: make false analogy or misanalysis, apply an incomplete rule or exploit redundancy, overlook co-occurrence restrictions, overuse monitor (hypercorrection), overgeneralize or simplify a system.

In the research there seem to be differences in the terminology regarding errors. For example, in his SLA framework Major (2001) divides the phonological errors into transfer and developmental ones, of which the former decrease and the latter increase over time, which means that among the intermediate and advanced learners who are the subjects of our study TL-based errors are expected to be more abundant than ones originating from L1 transfer. On the whole, interlingual errors appear to be referred to as transfer, systemic or global errors, while intralingual errors are called target-language based or developmental.

For the sake of consistency with the previous study (Nowacka, 2016) we have decided to refer to the errors as global and local ones although as argued above we are aware that this terminology might not be the most-widely used in second language acquisition research.

2. Method

2.1. Aims

In this study, we have targeted at finding categories, i.e. patterns or rules concerning letter-to-sound relations, that are not respected in the subjects' performance and recognition of an individual word and should be explicitly discussed and practised in phonetics courses. We intend to uncover the types of errors that are the most frequent in our respondents' production and recognition of words, whether avoidable globalised, 'either-or' or true local ones, as classified by Porzuczek (2015). Our final aim is to bring to light similarities and differences between speakers of different L1s, both in production and perception.

2.2. Instruments and administration

For the purpose of this study we have designed a two-task test, task one on production and task two on recognition, altogether 62 lexical items taken from Sobkowiak's (1996: 294) *Words Commonly Mispronounced* and more precisely from Porzuczek's selection of the first 373 words of the abovementioned Sobkowiak's list. Each task included 31 lexical items of Sobkowiak's (1996)

words commonly mispronounced – see Nowacka (2016) for the choice of lexical items and the questionnaire.

When it comes to test administration, there were two periods of recording sessions: the first in October 2015 with mainly Polish respondents and some Erasmus students of English at two universities; and the second in February 2016: with non-Polish respondents, on mostly science courses, who represented thirteen nationalities, after which we selected the five most numerous populations. One of the aims of the study was to examine the phonetic know-how of the freshmen of the English course before they undergo phonetic instruction. At the time of the data collection, Polish and non-Polish respondents had undergone no or hardly any prior phonetic training.

In the production task our respondents were asked to read thirty-one items together with the corresponding number and to record them. Their enunciation was then rated by the author of the text herself and classified as correct if it belonged to major standard Englishes within the scope of my knowledge. Thus, for example the renditions of *chair* as /tʃeə/ or /tʃe(ə)r/ were regarded as correct as opposed to an erroneous form /tʃɜː/ or for the word *author* both /ˈɔːθə/ and /ˈɑːθə/ were assigned a positive mark but not the mispronunciation /ˈəʊθə/. We are fully aware that it would have been much more reliable if a greater number of judges, preferably encompassing both native and non-native varieties of English, had been involved in this evaluation.

In the recognition task, they were exposed to a recording of thirty-one items, which they could see in the test written in standard spelling. Each item was pronounced twice, they heard the two pronunciation versions in a random order one version contained standard British English pronunciation of the word while the other was a deviant Polglish mispronunciation based on Sobkowiak's (1996) transcription. The subjects were required to point to the correct rendition of each item, by circling the letter A or B. Each item was repeated twice.

2.3. Subjects

There were altogether 291 participants in the study, of which 238 completed both production and recognition tasks (see Table 1).

Table 1. Bio-data

	Task 1: PRODUCTION		**Task 2: RECOGNITION**	
	n	%	*n*	%
Total	**238**	100	**291**	100
UNIVERSITY				
non-public university	129	54.2	164	56.4
public university	109	45.8	127	43.6
SYSTEM				
daily	174	73.1	217	74.6

	Task 1: PRODUCTION		**Task 2: RECOGNITION**	
extramural	39	16.4	46	15.8
postgraduate	25	10.5	28	9.6
SEX				
female	106	44.5	169	58.1
male	66	27.7	98	33.7
no data	66	27.7	24	8.2
NATIONALITY				
Polish	150	63.0	193	66.3
Ukrainian	36	15.1	38	13.1
Kazakh	27	11.3	27	9.3
Turkish	13	5.5	13	4.5
Tajik	7	2.9	10	3.4
Malaysian	5	2.1	10	3.4
FACULTY			*n*	%
art (English Language) (n=143)			175	73
science			63	26

The data was collected at two universities in Rzeszów, one private: the University of Information Technology and Management (164) and the other one public (127): the University of Rzeszow. There is a slight predominance of private (56.4%) over public (43.6%) students in number. In addition, they were mainly on daily (217), but also extramural (46) and postgraduate courses (28); mostly female (169) with one-third of males (98), no data was provided in 24 cases; predominantly students of Arts (175) but also Science (63).

As regards nationality, the Polish group was the most numerous (193 respondents), constituting 66.3%, then there were, in descending order: Ukrainians (38 – 13.1%), Kazakhs (27 – 9.3%), Turkish (13 – 4.5%), Tajiks (10 – 3.4%) and Malaysians (10 – 3.4%).

3. Results and discussion

3.1. Production

On the basis of lowest scoring results (including 5%–50% of correct renditions) in the ranking of 31 lexical items, representing 27 phonotactic patterns we observe that in the respondents' performance both local and global errors are present (see Table 2).

Words whose mispronunciations belong to **true local errors** such as *dough*, *above*, *southern*, *knowledge* and *area* together with one example of **either-or local errors**, that is *ancient*, obtain low scores and should be included in the core of pronunciation instruction because they are frequently mispronounced by speakers of six different L1s.

In addition, here we can also find seven rarely applied patterns, which belong to **avoidable globalised errors**, i.e. mute consonant letters (*comb* – whose frequency of use is rather low as represented by Band 5), 3 letter-to-sound vocalic rules (*old*, *layer* and *world*), 2 vowel reduction rules (*accurate*, *surface*) and a class of so-called isolated errors,[4] encompassing 15 patterns represented by *thousand* and *pronounce*.

What this finding implies is that pronunciation training should encompass explicit instruction of some spelling conventions for the benefit of learners' pronunciation. It is thus confirmed that learn-by-rote categories from Porzuczek's (2015) division include 1 either-or local error: unpredictable pronunciation of single vowel letters (*ancient*), and 3 true local errors: words with unpredictable pronunciation (*southern*, *knowledge*, *abroad* and *says)*, the sequence <-ough> (*dough*) and the letter 'o,' ('<o> → □/ɒ/ - /ʌ/ - /əʊ/ - (/u:/) - (/ʊ/)') (*above*).

These results allow us to advocate the explicit instruction of some productive phonotactic rules regarding the globalised errors that our respondents have made. This group includes vowel reduction rules in stress-adjacent and stress-following syllables (*surface*) and suffixes, e.g. <-ous>, <-age>, <-ate> (*accurate*); the ambiguous letter <o> as a whole, leading to different phonetic shapes, depending on the context, for example: '<-old> → /əʊld/; <oll> → /əʊl/ but (*doll*)' as in *old* and also the letter <o> in local errors being a part of the category 'words with unpredictable pronunciation' such as *southern*, *knowledge* and *abroad*, the occurrence of silent letters (*comb*) – the top word mispronounced by 95% of the respondents, but also a vocalic rule concerning NURSE, i.e. "stressed preconsonantal or word-final <wor>, <ur>, <ir>, <er> →/ɜ:/; <earC> → /ɜ:/ if C is not an inflectional ending (but *beard*)," and 'isolated errors' (unfamiliar to slightly more than a half of our respondents).

To our surprise the lexical item *area*, constituting a local error category of 'unpredictable word stress' scored higher than we expected (50% of correct renditions), although we know from other studies (Waniek-Klimczak, 2015) that 'stress placement' is usually problematic.

Table 2. Ranking of results for word production: task 1 (word reading)

No.	Phonotactic pattern:	FB**	Lexical item	%
1.	mute consonant letters (T.26)	5	*comb*	5%
2.	<-old> → /əʊld/; <oll> → /əʊl/ but (*doll*) (T.22)	*7*	*old*	8%
3.	<-aiC>, <-ay> → /eɪ/ (T.21)	*6*	*layer*	15%
4.	unpredictable pronunciation of single vowel letters (T.5)	*6*	*ancient*	20%

[4] Porzuczek (2015:186) notes that isolated errors "can be avoided if general spelling-to-sound rules are observed, even though the actual pronunciation is not always predictable."

** FB stands for frequency band.

No.	Phonotactic pattern:	FB**	Lexical item	%
5.	unpredictable <-ough> (T.3)	5	*dough*	25%
6.	Reduce <-ous>, <-age>, and <-ate> in nouns and adjectives (T.14)	*6*	*accurate*	29%
7.	<o> → /ɒ/ - /ʌ/ - /əʊ/ - (/uː/) - (/ʊ/) (T.4)	*7*	*above*	32%
8.	Reduce the vowel in stress-adjacent syllables and in syllables following the stressed one to /ə/ or /ɪ/. (T.13)	*7*	*surface*	32%
9.	words with unpredictable pronunciation (T.1)	*6* *7* *6* *6*	*southern* (22%), *knowledge* (33%), *says* (43%), *abroad* (44%)	35.5%
10.	isolated errors (T.27)	*6* *6*	*thousand (42%),* *pronounce (46%)*	44%
11.	Stressed preconsonantal or word-final <wor>, <ur>, <ir>, <er> →/ɜː/; <earC> → /ɜː/ if C is not an inflectional ending (but *beard*). (T.17)	*7*	*world*	47%
12.	unpredictable word stress (T.2)	*7*	*area*	50%
13.	<ou> → /aʊ/ - (/uː/) – (/ʌ/). <ou> ≠ /əʊ/ (T.8)	*6*	*youth*	51%
14.	<aw> → /ɔː/ (T.19)	*7*	*draw*	53%
15.	<-ought>, <-aught> → /ɔːt/ (but *drought*) (T.18)	5	*taught*	55%
16.	<(s)waC-> → /(s)wɒC/; <(s)quaC-> → /(s)kwɒC/; <war(C)> → /wɔː(C)/ (T.23)	*6*	*ward*	56%
17.	<air> → /eə/ (T.20)	*6*	*aircraft*	60%
18.	problems with voicing (T.11)	*7*	*basic*	60%
19.	<i> ≠ /iː/ (T.24)	*6*	*pitch*	63%
20.	predictable consonant voicing (T.25)	*7*	*pressure*	68%
21.	<ow> → /aʊ/ - /əʊ/ (T.7)	*6*	*bowl*	71%
22.	<ea> → /iː/ - /e/ - (/eɪ/) (T.6)	*6*	*breathe*	78%
23.	Never stress the adjectival *–able/-ible* suffix. Reduce it to /-əbl/ instead. (T.15)	*7*	*available*	84%
24.	If unstressed <-er>, <-our> → /ə/; <-ey> → /ɪ/ (T.16)	5	*donkey*	84%
25.	<g> → /g/ - /dʒ/ before <e>, <i>, <y> (T.12)	*6*	*target*	87%
26.	<ear> → /ɪə/ - /eə/ (T.10)	*6*	*ear*	88%
27.	<au> → /ɔː/ - (/ɒ/). <au> ≠ /əʊ/, /aʊ/. (T.9)	*7*	*because*	90%

3.2. Recognition

In general, the results in the recognition task are higher than those for production, which results from this being an easier type of task (see Table 3).

Table 3. Ranking of results for word recognition: task 2

No.	Phonotactic pattern:	FB	Lexical item	%
1.	<air> → /eə/ (T.20)	6	*chair*	22%
2.	problems with voicing (T.11)	7	*increase (v.)*	22%
3.	<ow> →□/aʊ/ - /əʊ/ (T.7)	5	*owl*	31%
4.	unpredictable pronunciation of single vowel letters (T.5)	5	*pint (22%), angel (61%)*	41.5%
5.	<-old> → /əʊld/; <oll> → /əʊl/ but (*doll*) (T.22)	6	*cold*	42%
6.	words with unpredictable pronunciation (T.1)	5	*failure* (31%), *colonel* (36%), *don't* (59%)	42.5%
7.	<(s)waC-> → /(s)wɒC/; <(s)quaC-> → /(s)kwɒC/; <war(C)> → /wɔ:(C)/ (T.23)	6	*wander*	45%
8.	<ou> →□/aʊ/ - (/u:/) – (/ʌ/). <ou> ≠ /əʊ/ (T.8)	6	*wound* (*n.* injury)	49%
9.	<-aiC>, <-ay> → /eɪ/ (T.21)	6	*layer*	57%
10.	mute consonant letters (T.26)	5 6	*hymn (48%), muscle* (76%)	62%
11.	stressed preconsonantal or word-final <wor>, <ur>, <ir>, <er> →/ɜ:/; <earC> → /ɜ:/ if C is not an inflectional ending (but *beard*). (T.17)	6	*worth*	64%
12.	predictable consonant voicing (T.25)	7	*though*	67%
13.	<au> →□/ɔ:/ - (/ɒ/). <au> ≠ /əʊ/, /aʊ/. (T.9)	7	*author*	68%
14.	reduce <-ous>, <-age>, and <-ate> in nouns and adjectives (T.14)	6	*enormous*	70%
15.	<g> →□/ɡ/ - /dʒ/ before <e>, <i>, <y> (T.12)	6	*gear*	71%
16.	<ea> → /i:/ - /e/ - (/eɪ/) (T.6)	5	*sweat*	71%
17.	Never stress the adjectival *–able/-ible* suffix. Reduce it to /-əbl/ instead. (T.15)	6	*capable*	72%
18.	unpredictable word stress (T.2)	7	*develop*	73%
19.	<-ought>, <-aught> → /ɔ:t/ (but *drought*) (T.18)	6	*ought*	73%
20.	<o> →□/ɒ/ - /ʌ/ - /əʊ/ - (/u:/) - (/ʊ/) (T.4)	7	*company*	76%
21.	isolated errors (T.27)	6	*variety*	79%
22.	unpredictable <-ough> (T.3)	5	*through*	80%
23.	Reduce the vowel in stress-adjacent syllables and in syllables following the stressed one to /ə/ or /ɪ/.	7	*certain*	81%
24.	<ear> →□/ɪə/ - /eə/ (T.10)	6	*tear* (*n.* eye water)	86%
25.	<i> ≠ /i:/ (T.24)	7	*picture*	88%
26.	<aw> → /ɔ:/ (T.19)	7	*law*	90%
27.	If unstressed <-er>, <-our> → /ə/; <-ey> → /ɪ/ (T.16)	6	*monkey*	95%

The lexical items that obtained the lowest scores, i.e. under 60%, are as follows: *chair*, *increase* (v.), *owl*, *pint*, *angel*, *cold*, *failure*, *colonel*, *don't, wander, wound* and *layer*. In this grouping we can see the same number of either-or local errors

(marked pale grey) and global ones (on a white background), and only one type of true local error (marked dark grey).

Five of the examined words include a problematic letter ‘o,’ leading to a wide array of vocalic qualities. Two of these lexical items fall into ‘words with unpredictable pronunciation’ (*colonel* and *don’t*) and others represent three letter-to-sound patterns such as: ‘<ow> → /aʊ/ - /əʊ/’ (*owl*), ‘<ou> → /aʊ/ - (/u:/) – (/ʌ/). <ou> ≠ /əʊ/’ (*wound* - n.), and ‘<-old> → /əʊld/; <oll> → /əʊl/ but (*doll*)’ (*cold*).

This finding agrees with Collins and Mees (2008: 112) and Zając (2015) who note that of all the letters, the letter <o> is associated with most pronunciation irregularities. Moreover, in Bell’s (2009) classification of 69 spellings, which have more than one sound, the most numerous group concerns graphemes with the letter ‘o’ (15), i.e.

> **o**: on - *o*nly, *o*nce, *o*ther; **-o**: go – d*o*; **oa**: road - br*oa*d; **o-e**: bone - d*o*n*e*, g*o*n*e*; **-oes**: toes - d*oe*s, sh*oe*s; **-oll**: roll - d*oll*; **-omb**: combat - bomb, c*omb*, t*omb*; **oo**: boot - f*oo*t, fl*oo*d; **-ot**: parrot - dep*ot*; **our**: sour - f*our*, j*our*ney; **ou**: sound - s*ou*p, c*ou*ple; **-ough**: bough - thr*ough*, r*ough*, tr*ough*; **ought**: bought – dr*ought*; **oul**: should - sh*oul*der, m*oul*d; **ow**: how - l*ow*.

Bell (2010a) observes that

> [t]he greatest difficulties in learning to read English are posed by irregular use of the letters **e** and **o**. They have variable sounds on their own (… on - ***o***nly, ***o***nce, ***o***ther, wh***o***) and in combinations with other letters: … b**o**n**e** – d***o***n***e***, g***o***n***e***; s**ou**nd – s***ou***p, s***ou***thern, sh***ou***lder; f**oo**d – fl***oo***d, g***oo***d; h**ow** – l***ow***. In all, 69 English spellings have more than one pronunciation and make around 2,000 relatively common words tricky for beginning readers, but the 205 with **o** … impede reading progress most of all, because they have different pronunciation in *very high frequency* words.

Bell (2010a) makes a comment that the letter o and combinations with o pose a problem for beginning readers because of their different and sometimes overlapping pronunciations. She lists the main sound for each grapheme with <o> together with other likely renditions, which we summarize here: <o> as in *on* (in an onset and nucleus position) usually stands for LOT but also for: GOAT (*only*), STRUT (*other*), GOOSE (*tomb*) and FOOT (*woman*), while as a coda the letter <o> represents GOAT (*go*) and GOOSE (*do*) only; <ou> is pronounced as MOUTH (*out*) but also as GOAT (*mould*), STRUT (*couple*), GOOSE (*group*) and FOOT (*could*); two syllable words with the letters <o-e> are rendered as GOAT (*bone*) as well as STRUT (*done*) and GOOSE (*move*); <oo> leads to GOOSE (*food*), GOAT (*brooch*) and STRUT (*flood*) and <ow> sounds as MOUTH (*now*) or GOAT (*slow*).

3.3. Similarities and differences between learners with different L1s

3.3.1. Production

No statistically significant differences have been observed in the pronunciation of 10 out of 31 lexical items among learners with different L1s (α=0,05). We present 5 phonotactic patterns in *youth*, *accurate*, *ancient*, *comb* and *southern* which are not applied by the majority (results around 50% and under with the exception of Malaysians in *youth*) (see Figure 1).

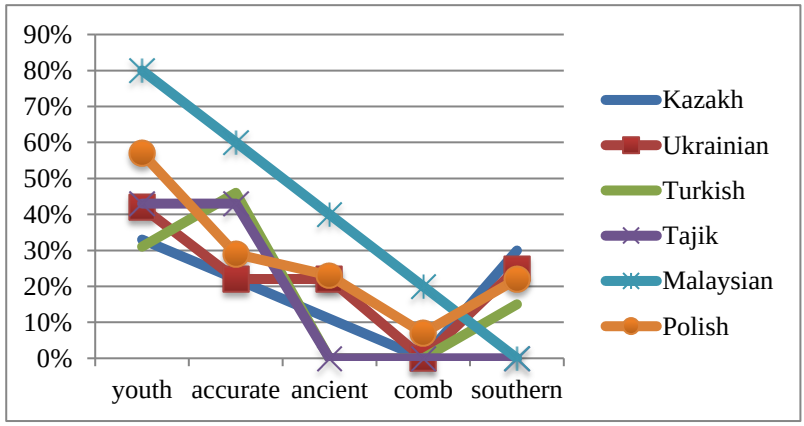

Figure 1. Statistically insignificant differences: similarities among learners with different L1s (α=0,05) – lexical items incorrectly pronounced by the majority of respondents

The patterns causing problems to the majority of respondents, regardless of their L1, are represented by: two avoidable globalised errors, i.e. 'reduce <-ous>, <-age>, and <-ate> in nouns and adjectives (T.14)' - *accurate* (p=.28845) and 'mute consonant letters (T.26)' - *comb* (p=.13778); two 'either-or' local errors, i.e. 'unpredictable pronunciation of single vowel letters (T.5)' - *ancient* (p=.12265) and '<ou> → /aʊ/ - (/u:/) – (/ʌ/). <ou> ≠ /əʊ/' (T.8) – *youth* (p=.05082); and one local error of 'words with unpredictable pronunciation (T.1)' - *southern* (p=.43551).

It needs to be added that when it comes to the frequency of occurrence of the above-mentioned words, four out of five lexical items belong to band 6, except for *comb* whose frequency is lower (band 5).

What has been confirmed by these results, and what can have practical implications for pronunciation teaching to learners with six L1s, i.e. Kazakh, Ukrainian, Turkish, Tajik, Malay and Polish, is that all of them would benefit from memorization of lexical items presented under the name of words with unpredictable pronunciation (Porzuczek 2015: 173); and explicit instruction on such issues as: words with mute consonant letters, unpredictable pronunciation of single vowel letters (Porzuczek 2015: 177), the focus on the ambiguous character

of the letter <o> and especially its diagraphs and the stress reduction rule concerning suffixes <-ous>, <-age>, and <-ate> in nouns and adjectives (Porzuczek 2015: 182).

However, in the articulation of 21 lexical items representing 18 phonotactic patterns statistically significant differences have been observed among speakers of different L1s (α=0,05) (see Table 4).

Table 4. Statistically significant differences in production among learners with different L1s

No.	Category	Lexical item	p
	LOCAL ERRORS:		
1.	words with unpredictable pronunciation (T.1)	*knowledge*	p=.00676**
		says	p=.00083***
		abroad	p=.00025***
2.	unpredictable word stress (T.2)	*area*	p=.00173**
3.	unpredictable <-ough> (T.3)	*dough*	p=.00011***
4.	<o> → /ɒ/ - /ʌ/ - /əʊ/ - (/u:/) - (/ʊ/) (T.4)	*above*	p=.00007***
	'EITHER-OR' LOCAL ERRORS		
5.	<ea> → /i:/ - /e/ - (/eɪ/) (T.6)	*breathe*	p=.00002***
6.	<au> → /ɔ:/ - (/ɒ/). <au> ≠ /əʊ/, /aʊ/. (T.9)	*because*	p=.00003***
7.	problems with voicing (T.11)	*basic*	p=.00015***
8.	<g> → /g/ - /dʒ/ before <e>, <i>, <y> (T.12)	*target*	p=.03347*
	GLOBALISED (AVOIDABLE) ERRORS		
9.	Reduce the vowel in stress-adjacent syllables and in syllables following the stressed one to /ə/ or /ɪ/. (T.13)	*surface*	p=.01644*
10.	Stressed preconsonantal or word-final <wor>, <ur>, <ir>, <er> →/ɜ:/; <earC> → /ɜ:/ if C is not an inflectional ending (but *beard*). (T.17)	*world*	p=.01560*
11.	<-ought>, <-aught> → /ɔ:t/ (but *drought*) (T.18)	*taught*	p=.02895*
12.	<aw> → /ɔ:/ (T.19)	*draw*	p=.00287**
13.	<air> → /eə/ (T.20)	*aircraft*	p=.00000***
14.	<-aiC>, <-ay> → /eɪ/ (T.21)	*layer*	p=0.0000***
15.	<-old> → /əʊld/; <oll> → /əʊl/ but (*doll*) (T.22)	*old*	p=.01566*
16.	<(s)waC-> → /(s)wɒC/; <(s)quaC-> → /(s)kwɒC/; <war(C)> → /wɔ:(C)/ (T.23)	*ward*	p=.02350*
17.	<i> ≠ /i:/ (T.24)	*pitch*	p=.02500*
18.	isolated errors (T.27)	*pronounce*	p=.00000***
		thousand	p=.00128**

The variety of differences between the six nationalities in the rendition of these words does not allow us to make generalisations. Though Malaysians as second

language learners of English are the most accurate when it comes to the pronunciation of a majority of these words (see Figure 2).

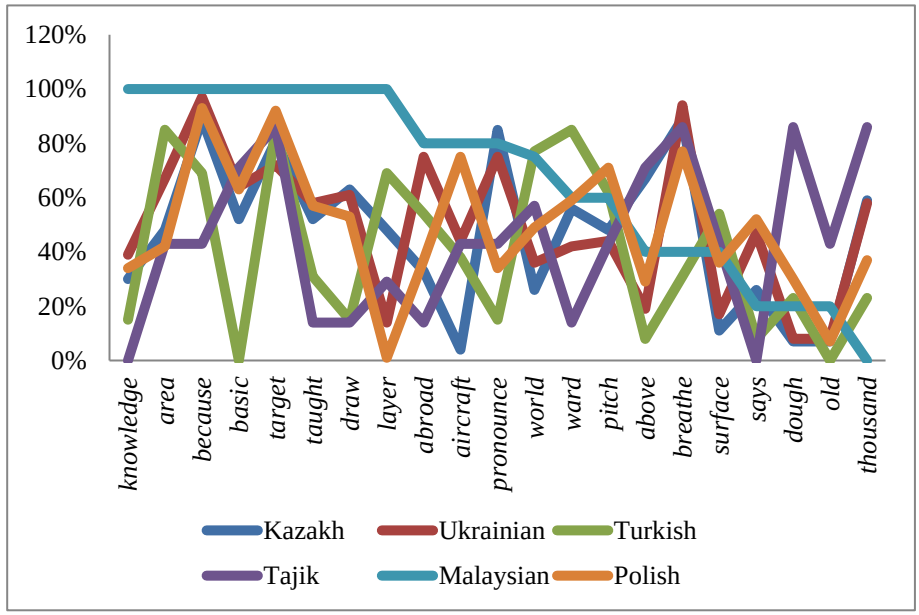

Figure 2. Statistically significant differences in production of lexical items among learners with different L1s

Our participants differ in the accuracy of pronunciation with respect to these lexical items. Each item has to be examined individually in search of similarities and differences between different L1 speakers. For example, Tajiks do not err on *dough* (86%) but other nationalities do, and Turkish students have no problem with the right quality of the vowel in *ward* but learners with the other 5 L1s show lower levels of correctness.

3.3.2. Recognition

In the recognition task there are no statistically significant differences among different L1 learners in twelve out of thirty-one cases. Three of twelve lexical items were not familiar in their pronunciation to all learners, regardless of their L1 (under 50% results), i.e. *wander*, *pint* and *increase* (see Figure 3).

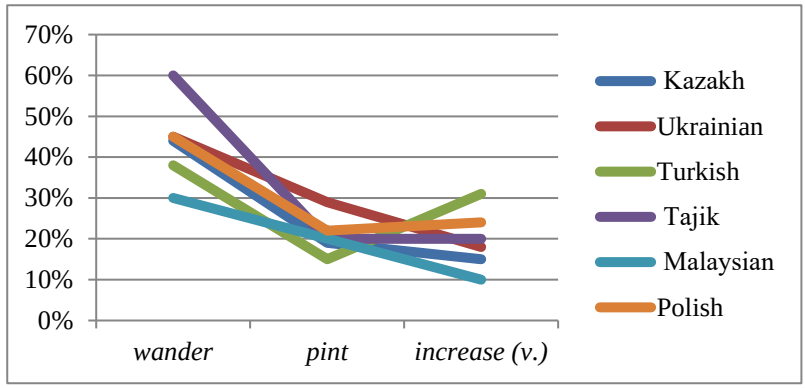

Figure 3. Statistically insignificant differences: similarities among learners with different L1s – lexical items incorrectly recognised by the majority of respondents

This group encompasses: two 'either-or' local errors such as 'problems with voicing (T.11)' - *increase (v.)* (p=.68301), 'unpredictable pronunciation of single vowel letters (T.5)' - *pint* (p=.90008) and one avoidable globalised error: '<(s)waC-> → /(s)wɒC/; <(s)quaC-> → /(s)kwɒC/; <war(C)> → /wɔ:(C)/ (T.23)' - *wander* (p=.84388). Thus, it would be useful to find time in a practical phonetics course to make learners aware of these issues that are not known to learners with these 6 different L1s.

The remaining 19 words, constituting 16 phonotactic categories, (see Table 5), show statistically significant differences for recognition in learners of different L1 backgrounds. As was the case with production, each lexical item and the corresponding phonotactic pattern has to be examined individually due to the abundance of data, e.g. we observe similarity in the correct recognition of the pronunciation of *colonel* by Malaysians and Poles (50%) differing from other nationalities for whom the range is up to 20% (see Figure 4).

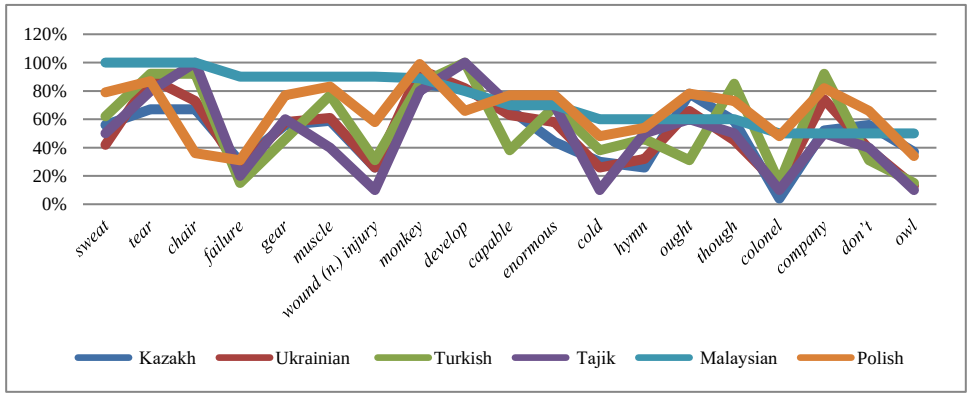

Figure 4. Statistically significant differences in recognition of lexical items among learners with different L1s

Table 5. Statistically significant differences in recognition among learners with different L1s

No.	Category	Lexical item:	p
	LOCAL ERRORS		
1.	‘words with unpredictable pronunciation (T.1)’	*colonel*	p=.00000
		don’t	p=.00562
		failure	p=.00159
2.	‘unpredictable word stress (T.2)’	*develop*	p=.00853
3.	‘<o> → /ɒ/ - /ʌ/ - /əʊ/ - (/u:/) - (/ʊ/) (T.4)’	*company*	p=.00056
	‘EITHER-OR’ LOCAL ERRORS		
4.	‘<ea> → /i:/ - /e/ - (/eɪ/) (T.6)’	*sweat*	p=.00001
5.	‘<ow> → /aʊ/ - /əʊ/ (T.7)’	*owl*	p=.03044
6.	‘<ou> → /aʊ/ - (/u:/) – (/ʌ/). <ou> ≠ /əʊ/ (T.8)’	*wound (n.)*	p=.00000
7.	‘<ear> → /ɪə/ - /eə/ (T.10)’	*tear (n.)*	p=.04695
8.	‘<g> → /g/ - /dʒ/ before <e>, <i>, <y> (T.12)’	*gear*	p=.00634
	GLOBALISED (AVOIDABLE) ERRORS		
9.	‘Reduce <-ous>, <-age>, and <-ate> in nouns and adjectives (T.14)’	*enormous*	p=.00930
10.	‘Never stress the adjectival –able/-ible suffix. Reduce it to /-əbl/ instead. (T.15)’	*capable*	p=.03797
11.	‘If unstressed <-er>, <-our> → /ə/; <-ey> → /ɪ/ (T.16)’	*monkey*	p=.00046
12.	‘<-ought>, <-aught> → /ɔ:t/ (but *drought*) (T.18)’	*ought*	p=.00356
13.	<air> → /eə/ (T.20)	*chair*	p=.00305
14.	‘<-old> → /əʊld/; <oll> → /əʊl/ but (doll) (T.22)’	*cold*	p=.01479
15.	‘predictable consonant voicing (T.25)’	*though*	p=.00602
16.	‘mute consonant letters (T.26)’	*hymn*	p=.02435
		muscle	p=.00041

4. Conclusions

The results of the study confirm the necessity for explicit instruction on the regularity rather than irregularity of English spelling in order to eradicate globalised and ‘either-or’ pronunciation errors in the speech of university students with six different L1s. The avoidable globalised errors which have turned out to be the most numerous in the production task include such areas of English phonotactics as: the letters <-old> and <oll>, ‘mute consonant letters’ (all 6 L1s), ‘isolated errors’ and two categories related to the reduction of unstressed syllables: ‘reduce the vowel in stress-adjacent syllables and in syllables following the

stressed one to /ə/ or /ɪ/,' 'reduce <-ous>, <-age>, and <-ate> in nouns and adjectives' (all 6 L1s).

Once introducing spelling-to-sound relations becomes a routine procedure in pronunciation training, the strain on the part of the students of memorizing a list of local errors, phonetically challenging pronunciation exceptions, will be reduced to the absolute minimum, comprising such aspects as: the ambiguous letter <o> (all 6 L1s), 'unpredictable <-ough>,' words with unpredictable pronunciation (all 6 L1s), unpredictable pronunciation of single vowel letters (all 6 L1s) and unpredictable stress placement.

It is believed that the outcome of our research makes it easier for teachers of phonetics to decide which graphophonemic patterns should be explicitly taught in phonetic instruction. We also hope that learners' production of some phonetically challenging items will improve if they make an attempt at memorizing some spelling guidelines, which we have ranked according to their needs.

We feel obliged to admit that there are some limitations to the methodology applied. It would have been more appropriate to implement a control group in the experiment, which would have involved running the test not only in a population of non-native students but also with native speakers of English. Another issue that requires improvement is the size of the populations selected for such an analysis. The groups corresponding to nationalities should have been of more or less equal number of respondents – we could have reduced the overwhelming number of Polish respondents to make this group comparable to others.

References

Bell, Masha. 2009. Reading problems. *English Spelling Problems blog.* [Online] Available from: http://englishspellingproblems.blogspot.com/2015/01/the-english-spelling-system.html?view=sidebar [Accessed: 19th February 2017].

Bell, Masha. 2010a. E, ea, ei, ie, a and o reading problems. *English Spelling Problems blog.* [Online] Available from: http://improvingenglishspelling.blogspot.com/2010/10/e-ea-ei-ie-and-reading-problems.html [Accessed: 19th February 2017].

Bell, Masha. 2010b. O-e / oa / ol / o / ow / ou ... . *English Spelling Problems blog.* [Online] Available from: http://improvingenglishspelling.blogspot.com/2010/08/o-e-oa-ol-o-ow-ou.html [Accessed: 19th February 2017].

Bell, Masha. 2015. The English Spelling System. *English Spelling Problems blog.* [Online] Available from: http://englishspellingproblems.blogspot.com/2015/01/the-english-spelling-system.html [Accessed: 19th February 2017].

Bryła-Cruz, Agnieszka. 2013. English. Irish and Scottish listeners' perception of Polish-accented English: a comparative study. [unpublished PhD dissertation] Lublin: Uniwersytet Marii Curie-Skłodowskiej.

Carney, Edward. 1994. *A Survey of English Spelling*. London: Routledge.

Ellis, Nick C. and S. Cataldo. 1990. The role of spelling in learning to read, *Language and Education:* 1–28.

English Spelling Society. 2017a. Irregularities. The English Spelling System. [Online] Available from: http://spellingsociety.org/irregularities-of-english-spelling#/page/1 [Accessed: 10th February 2017].

English Spelling Society. 2017b. Our aims. [Online] Available from: http://spellingsociety.org/our-aims [Accessed: 19th February 2017].

James, Carl. 1998. *Errors in language learning and use: Exploring error analysis*. London: Longman.

Jassem, Wiktor. 1971. *Podręcznik wymowy angielskiej*. Warszawa: PWN.

Jones, Daniel. 1932. *An Outline of English Phonetics*. Cambridge: W. Heffer & Sons Ltd.

Lewis, Jack Windsor. 2017. Blog 532 More on Weakforms (xx)nobody, no-one, nor, obvious(ly), on, our. *PhonetiBlog*. [Online] Available from: http://www.yek.me.uk/Blog.html#blog532 [Accessed: 15th February 2017].

Majer, Jan. 2002. Sick or Seek? Pedagogical Phonology in Teacher Training. In: Ewa Waniek-Klimczak and P. J. Melia (eds.) *Accents and Speech in Teaching English Phonetics and Phonology*, 153–176. Berlin: Peter Lang.

Major, Roy C. 2001. *Foreign accent: The ontogeny and phylogeny of second language phonology.* Mahwah: NJ: Lawrence Erlbaum Associates.

[no name] 2016. Key to Frequency, *OED Oxford English Dictionary: The definite record of the English language*. Oxford: Oxford University Press. [Online] Available from: http://www.oed.com [Accessed: 11th May 2016].

Nowacka, Marta, Piotr Cymbalista and Grzegorz A. Kleparski. 2011. *Sally Meets Harry: Practical Primer to English Pronunciation and Spelling*. Rzeszów: Wydawnictwo Uniwersytetu Rzeszowskiego.

Nowacka, Marta. 2016. English spelling among the top priorities in pronunciation teaching: Polglish local versus global(ised) errors in the production and perception of words commonly mispronounced. In: Ewa Waniek-Klimczak, Iwona Witczak-Plisiecka, Jan Majer (eds.), *Research in Language*, vol. 14 (2), 123–148. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Pęzik, Piotr and Magdalena Zając. 2012. Developing a corpus-based index of commonly mispronounced words. Book of Abstracts: Accents 2012. Paper presented at *Accents 2012 – VIth International Conference on Native and Non-native Accents of English*, Łódź, Poland, 6-8 December, 2012: 26, [Online] Available from: http://filolog.uni.lodz.pl/accents/Abstracts-2012.pdf [Accessed: 11th May 2016].

Porzuczek, Andrzej. 2015. Handling Global and Local English Pronunciation Errors. In: Ewa Waniek-Klimczak and Mirosław Pawlak (eds.), *Teaching and Researching the Pronunciation of English: Studies in Honour of Włodzimierz Sobkowiak, Second Language Learning and Teaching*, 169–187. London: Springer.

Richards, Jack C. (ed.) 1974. *Error analysis: Perspectives on Second Language Acquisition*. London: Longman.

Scheuer, Sylwia. 1998. Interference-motivated pronunciation errors in Polish students of English. A corpus-based study. [unpublished PhD dissertation] Poznań: Adam Mickiewicz University.

Sobkowiak, Włodzimierz 1996. *English Phonetics For Poles*. Poznań: Bene Nati.

Szpyra-Kozłowska, Jolanta. 2005. Intelligibility versus Polish accent in English. *Studia Phonetica Posnaniensia*. *An International Journal for Linguistic Phonetics*, vol. 7, 59–73. Poznań: Wydawnictwo Naukowe Poznań.

Szpyra-Kozłowska, Jolanta. 2013. On the irrelevance of sounds and prosody in foreign-accented speech. In: Ewa Waniek-Klimczak and Linda Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, 15–29. Berlin: Springer.

Szpyra-Kozłowska, Jolanta. 2015. *Pronunciation in EFL Instruction: A Research-Based Approach*. Bristol: Multilingual Matters.

Szpyra-Kozłowska, Jolanta and Sławomir Stasiak 2010. From focus on Sounds to Focus on Words in English Pronunciation Instruction. *Research in Language*, vol. 8 Doi: 10.1478/v10015-010-0012-7 [Online] Available from: http://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.hdl_11089_9575/c/v10015-010-0012-7.pdf [Accessed: 11th May 2016].

Szyszka, Magdalena. 2003. Efekty dydaktyki wymowyjęzyka angielskiego w oparciu o transkrypcję fonetyczną. In" Włodzimierz Sobkowiak and Ewa Waniek-Klimczak (eds.), *Dydaktyka fonetyki*

języka obcego II. Konferencja w Wąsoszach, 10-12.5.2002 [Phonetics in FLT 2, Wąsosze 10-12.5.2002], (=Zeszyty Naukowe Państwowej Wyższej Szkoły Zawodowej w Koninie). 2003, 145–149. Konin: Wydawnictwo PWSZ.

Upward, Christopher and George Davidson. 2011. *The History of English Spelling*. Chichester: Wiley-Blackwell.

Waniek-Klimczak, Ewa. 2015. Factors Affecting Word Stress Recognition by Advanced Polish Learners of English. In" Ewa Waniek-Klimczak and Mirosław Pawlak (eds.), *Teaching and Researching the Pronunciation of English: Studies in Honour of Włodzimierz Sobkowiak. Second Language Learning and* Teaching, 189–204. London: Springer.

Wells, John C. 2010. Non-spelling pronunciation, *Phonetic blog*. [Online] Available from: http://phonetic-blog.blogspot.com/2010/11/non-spelling-pronunciation.html [Accessed: 14th February 2017].

Wells, John C. 2011a. bilding a cubbard. *Phonetic blog*. [Online] Available from: http://phonetic-blog.blogspot.com/search?q=bilding [Accessed: 14th February 2017].

Wells, John C. 2011b. degreasing the galleon. *Phonetic blog*. [Online] Available from: http://phoneticblog.blogspot.com/search?q=galleon [Accessed: 14th February 2017].

Wells, John C. 2011c. Honorarium. *Phonetic blog*. [Online] Available from: http://phonetic-blog.blogspot.com/search?q=honorarium [Accessed: 14th February 2017].

Wells, John C. 2012. ideas. *Phonetic blog*. [Online] Available from: http://phoneticblog.blogspot.com/search?q=ideas [Accessed: 14th February 2017].

Zając, Magdalena. 2015. Compiling a Corpus-Based list of Words Commonly Mispronounced. In: Ewa Waniek-Klimczak and M. Pawlak (eds.) *Teaching and Researching the Pronunciation of English: Studies in Honour of Włodzimierz Sobkowiak, Second Language Learning and Teaching*, 153–167. London: Springer.

OLD ENGLISH SIMILE OF EQUALITY: THE HIGHEST DEGREE OF SIMILARITY

MARIANA OLENIAK
Vasyl` Stus Donetsk National University, Ukraine
maryana@mail.org

Abstract
This paper aims to provide an account of Old English similes of equality marked by the superlative degree of the adjective *gelic*. It deals with the structure and semantics of similes marked by the (*ge)/(on)licost* component, which, unlike in Modern English, being subjected to gradation, can show the highest degree of similarity between referents. The article presents the criteria for structural classification of the simile in question describing two major structural types, that employ nouns in the dative or nominative case, as well as its semantic interpretation from macro and micro levels of perspective. The paper examines every simile with the (*ge)/(on)licost* component found in Old English manuscripts belonging to various textual genres.

Keywords: Old English, simile, tenor, vehicle, salient feature

1. Introduction

Considering a fine line between metaphor and simile one must admit that the former becomes the focus of analyses much more often than the latter not only in linguistics, but also in cognitive psychology, philosophy, literary theory and criticism. Even less is simile investigated in the earliest periods of the English language. To comprehend the scale of difference in the amount of research devoted to these tropes one could consider google search data, that offer about 261000 results for a “metaphor in English”, 37000 results for a “simile in English” and only 2 results for a “simile in Old English”. Orval (1951) studies Old English simile in poetry and emphasizes their scarcity. Cavill (1999) comes to the same conclusion. To the best knowledge of the author the only diachronic investigation of the structure of similes in Middle English was conducted by Nevanlinna (1993) based on the Helsinki corpus. There have also been some sporadic cases of simile analyses as a part of syntactical investigation of Old English comparative clauses or the dative case functioning (Baker (2012), Fisher (1992), Gergel (2008), McLaughlin (1983), Merritt (2013), Mitchel (1985)) as well as in the literary works of certain periods or some authors’ works. Thus, Margolis (1957) draws a comparison between simile and related devises; Walker (2016), Beardsley (1981), Dawes (1998) study semantics of simile; Tyler (2006), highlights the poetics of the Old English literature providing some simile examples; Amodio (2014),

Stodnick (2012), Trilling (2012) analyse and compare Old English translations and adaptations of Latin texts some of them containing similes.

This study aims to provide an account of every simile in Old English that contains the superlative degree of the Old English adjective *gelic* from the point of view of their structure and semantics. The data discussed are taken from the Old English corpus of University of Toronto employing a continuous sampling method. The Old English corpus of the University of Toronto contains at least one copy of every surviving Old English text, thus all existing cases of the abovementioned type of simile have been analysed. The whole corpus of the Old English similes of equality containing the superlative degree of a copulative *lice* contains 31 examples.

2. The structure of the Old English simile with the (*ge)/(on)licost* component

Simile has been the tool of rhetoric for ages. It surely should be recognized as being universal, used by all speakers irrespective of their nationality, but manifesting itself differently in different languages. Given that many Old English texts are translations or interpretations from Latin and can only be assigned the date of their creation depending on the survived manuscript, the author considers the Anglo-Saxon simile not so much as a manifestation of a personal craft, being an individually inspired production of a genius, but as a collective, participatory, intellectual and cultural output echoing with current life, conventions, and complying with the current system of information exchange.

Old English employs certain linguistic mechanisms to create simile of equality among which there is the use of the superlative degree of a copulative *lice* in a variety of structural combinations, for example:

(1)
Foran æghwylc wæs, <stiðra> nægla gehwylc, style gelicost, hæþenes
‘Before each was <stiff> claw-nail, to steel most like, heathen's
handsporu <hilderinces>, egl', unheoru
banged up <worrior’s>, grip not gentle’

The terms open/closed simile are usually used in the context of simile semantics investigations (Margolis 1957; Beardsley 1981; Dawes 1998; Walker 2016) but taking into account the fact that any changes in semantics trigger corresponding changes of form, we will use these characteristics of simile as a useful criteria to organize all the structures of similes under consideration into a consistent paradigm. The criteria for our structural classification are: 1) the case of the noun used as a vehicle; 2) the amount of the simile components explicated in the surface structure: three (tenor, vehicle and a comparison marker) or all four (tenor, vehicle, comparison marker and a common salient feature shared by the tenor and the vehicle); 3) their grammatical expression; 4) their positioning.

Notwithstanding the fact that Old English was a highly-inflected language with a relatively free word order, we can observe a strong tendency to place the topic in a preposition in respect to the vehicle (1) though there are some instances of the contrary situation as in Beowulf:

(2)
… *him of eagum stod ligge gelicost leoht unfæger*.
'…from his eyes stood to a flame most like light horrible.'

The positioning of the comparison marker was not regulated by hard-and-fast rules: it can either precede (example 11) (33 % of examples) or follow the vehicle (example 1) (67 % of examples) but it never occurs in the initial position within the simile construction and it is always in contact position with the vehicle whilst other simile components can be either in contact (example 4) or in a distant position (example 3) relating to one another. The spelling of it is rather inconsistent and sometimes idiosyncratic since six different variants have been registered: *gelicost* – 21 examples (67,7%), *geliccost* – 2 examples (6,5%), *geliccast* – 1 example (3,2%), *gelicast* – 4 examples (13%), *gelicust* – 1 example (3,2%), *olnicost* – 1 example (3,2%) and *licost* – 1 example (3,2%). The position of the comparison marker defines the case of the noun or the pronoun used as a vehicle: if it follows the vehicle, the latter is used in the dative case (example 6), if it proceeds the vehicle, the latter is used in the nominative case (example 10). It is worth mentioning that the combination of the dative case and a superlative degree of *lice* was highly productive in the Old English simile formation: It constitutes 27 out of 31 examples, which is 90% of the sampling.

The whole scope of similes in question can be classified into two major structural types with the further subdivision: 1) Old English similes with the *gelicost* component employing a noun in the dative case as a vehicle; 2) Old English similes with the *gelicost* component employing a noun in the nominative case as a vehicle.

2.1. Old English similes with the *gelicost* component employing a noun in the dative case as a vehicle

The basic structure for the majority of Old English similes containing different grammatical forms of a copulative *lice* is N / Pron + N_{dat} +*gelicost*. Similes employing a noun in the dative case as a vehicle can start with either a tenor or a vehicle and, thus, can be subdivided into two subtypes.

2.1.1. The distinctive feature of the first structural subtype of the Old English similes with the gelicost component is that it starts with a tenor.

It can be either open (when it does not explicate a common feature shared by the topic and the vehicle of the simile: subtypes of 2.1.1.1.) or closed (when it explicates a common feature shared by the topic and the vehicle: subtypes of

2.1.1.2.). Being open, it consists of three elements: the first component can be either a noun or a pronoun representing the tenor of the simile, the second component is a noun in the dative case being the vehicle of it and the third one is the comparison marker. The empirical material gives the reason to divide this structural type into two subtypes. The first is the basic structure proper:

2.1.1.1. Open simile of the structure: N / Pron + N_{dat} +gelicost.

Open simile of the structure: N / Pron + N_{dat} +*gelicost* can be devided into two subtypes.

a) The first subtype represents the basic structure proper: N / Pron + N_{dat} +*gelicost*

(3)

Word spearcum fleah attre gelicost, þonne he ut þorhdraf: "Eala drihtenes þrym

'Words in sparks flew out to poison most like, when he exploded: "Alas the Lord's glory!"'

In the example (3) taken from "Christ and Satan", that, presumably, dates back to the 7-th century, "the words" are compared to "poison" with no salient feature explication. The distant position of the tenor and the vehicle in this structural type was optional.

b) The second subtype of this structure is extended by a noun phrase that specifies the vehicle but cannot be regarded as an explicit commonly shared feature since it is not semantically specific enough to derive this property: N + N_{dat} +*gelicost* +NP (Adj + N + clause)

(4)

Is seo eaggebyrd stearc ond hiwe stane gelicast, gladum gimme, þonne in goldfate

'Is that bird's eye piercing and of colour to stone most like, a bright gem,

smiþa orþoncum biseted weorþeð.

when by a skillful smiths' in beaten-out gold is set.'

This structural type is very similar to the first one, but the vehicle here is further specified by an extended noun phrase. Thus, the colour of the eye is not just compared to any stone, but a stone, which is a bright gem, and not just any bright gem, but the one that is skilfully put into gold. The colour in question is not mentioned though, so, notwithstanding an extended specification of the kind of the gem, the simile cannot be considered a closed one.

2.1.1.2. Closed simile of the structure: N / Pron + N_{dat} +gelicost.

If a simile of this structural type is closed, the composition of it is complicated by additional lexemes, phrases or clauses which are the verbalisation of the concept making the simile closed (further in the text they are typed in bold). Four subtypes of such a structure can be singled out.

a) N +**Adj** + N_{dat} +*gelicost*+clause

(5)
se engel becwom, windig and wynsum, wedere gelicost þonne hit on
*'…*angel descended, windy and pleasant, to weather most like, when in the summer season
sumeres tid sended weorðeð dropena drearung on dæges hwile, wearmlic wolcna scur.
drops of rains are sent falling during the day, and warm showers from the heavens.'

This structural type is based on the structure N / Pron + N_{dat} + *gelicost*; with the additional adjectives used between the tenor and the vehicle that explicate the commonly shared features, besides, the clause is considered to be a part of the simile since it specifies the vehicle. So, in the example, given above, an angel is compared to breezy summer weather on the grounds that both of them are windy and pleasant.

b) N +**V/VP** + N_{dat} +*gelicost*

The structural subtype in question is quite a productive one. What differentiates it from the previously described ones is that the salient feature, shared by the topic and the vehicle, is expressed either by a verb or a verb phrase and not by an adjective.

(6)
Landes frætwe gewitaþ under wolcnum winde geliccost.
'The lands will depart under the sky, to the wind most like.'

In this 9-th century example taken from "The Phoenix" the land is compared to wind since both can depart, though the character of the departure is not quite clear: whether or not the departure is fast or intangible or temporary, etc.

c) **clause**+N + N_{dat} +*gelicost*

(7)
"*Gewat þa ofer wægholm, winde gefysed, flota famiheals fugle gelicost…*"
'Went over the sea of waves, by the wind driven, the foamy-necked floater to a bird most like…'

In (7) the author of Beowulf compares the ship to a bird since they, if moved by the wind, float on the water in a very similar way. Thus, he specifies the simile with a clause explaining why a ship and a bird are very much alike.

d) N + $N_{dat/}$$Pron_{dat}$ + *gelicost* +**clause**

(8)
Het wæpen wera wexe gelicost on þam orlege eall formeltan ….
'(he) let the weapons to be to wax most like, completely melt in that fight …'

This structural subtype is similar to the previous one differing only in the positioning of the clause containing a commonly shared salient feature that is in

postposition in this case. The given example shows the comparison of the weapon to wax having the same property – being able to melt entirely.

2.1.2. The distinctive feature of the first structural subtype of the Old English similes with the gelicost component is that it starts with a vehicle.

The distinctive feature of this structural subtype of the Old English similes with the *gelicost* component is that it starts with a vehicle: N_{dat} +*gelicost* +clause / $N_{nom}P$:

(9)
….. wundum weorðan, wætre geliccost faran flode blod.
'…wound was made, to rain most like, flew floods of blood.'

Unlike all the other structural types, this one opens with a vehicle that is followed by the *gelicost* and the clause containing the tenor. Thus, the (9) shows the similarity between the flowing of blood and the rain, placing the dative case of the noun "rain" in preposition to the comparison marker and the tenor.

All the previously described structural types employed a noun in the dative case to mark the vehicle of the simile. The author will further proceed with the cases in which the vehicle of the simile is expressed by a noun in the nominative case. The following structural types display only closed similes.

2.2. Old English similes with the *gelicost* component employing a noun in the nominative case as a vehicle.

The distinctive feature of the second structural type of the Old English similes with the gelicost component is that it employs a noun in the nominative case as a vehicle. It can be subdivided into two subtypes depending on the grammatical function of a vehicle and a verbalized ground of the simile.

2.2.1. Vehicle is a separate noun further specified by a clause

In the first subtype of the structural type in question the vehicle is expressed by a separate noun further specified by a clause: N +**Adj**+ *onlicost* + $N_{nom}P$ (N_{nom}+clause).

(10)
Se fugel is on hiwe æghwæs ænlic, onlicost pean, wynnum geweaxen,
'The bird in colour is in every way unique, most like to a peacock, grown
þæs gewritu secgað.
with joys, of which writings speak'.

In (10) taken from "The Phoenix", the adjective *ænlic* 'unique' refers both to the colour of the phoenix and a peacock serving as a basis for their juxtaposition; the vehicle, that immediately goes after the comparison marker is extendedly

specified by a clause, that is why the uniqueness of the colour of the bird is not just compared to any peacock, but to the one grown with joys, of which writings speak.

2.2.2. Vehicle is the subject of the situation described by a clause

This structural subtype differs from the others in the way that the vehicle of the structure is not expressed by a separate noun further specified by a clause, but it is the subject of the situation described by a clause that explains the common characteristics of the juxtaposed concepts. Being the subject of a clause, it is used in the nominative case: N+**V**+*gelicost*+clause.

(11)
"& his æfterfolgeras feowertiene gear hit siþþan totugon & totæron
'and his fourteen successors and a year later took hold of them and tore
þæm gelicost þonne seo leo bringð his hungregum hwelpum hwæt to etanne…"
them most like to when the lion brings his hungry cubs something to eat….'

The first element of this structural type is a noun, representing the tenor, the second is a verb, which is a commonly shared feature by the tenor and the vehicle, and, joined by the *gelicost*, the clause contains the vehicle of the simile in the nominative case describing the way in which the salient feature is realized. So, in the provided example from Orosius (11), that belongs to the late Old English period, people are compared to lion cubs on the grounds of performing an analogous action – tearing the object apart.

The statistical data of the singled out structural types of the Old English simile containing a *gelicost* component is shown in the Table 1.

Table 1. The structural types of the Old English simile of equality containing a *gelicost* component

Type		Amount
2.1.	Old English similes with the *gelicost* component employing a noun in the dative case	
2.1.1.	N / Pron + N_{dat} +*gelicost*	**26**
2.1.1.1.	open simile:	
a)	N / Pron + N_{dat} +*gelicost*	6
b)	N + N_{dat} +*gelicost* +NP (Adj + N + clause)	1
2.1.1.2.	closed simile:	
a)	N +**Adj** + N_{dat} +*gelicost*+clause	1
b)	N +**V/VP** + N_{dat} +*gelicost*	9
c)	**clause**+N + N_{dat} +*gelicost*	1
d)	N + $N_{dat/}$Pron$_{dat}$ + *gelicost* +**clause**	8
2.1.2.	N_{dat} +gelicost+*clause / $N_{nom}P$*	**2**
2.2.	Old English similes with the *gelicost* component employing a noun in a nominative case	
2.2.1.	N +**Adj**+ *onlicost* + N_{nom}P (N_{nom}+clause)	**2**

Type	Amount
2.2.2. N+**V**+*gelicost*+clause	**1**
Total	**31**

The figures show a strong tendency to start simile with a tenor preceding a vehicle, expressed by a masculine or a neuter noun in the dative case, followed by the *gelicost* component. Since this positioning is true for the majority of the similes in question, it can be called a typical syntactic pattern of the Old English simile with the *gelicost* component. The employment of the nominative case vehicle was less frequent (3 cases out of 31) though representing 10% of the total sampling, it can be said to have been an alternative to the dative case simile structure being its grammatical synonym. A commonly shared salient feature, if explicitly marked, is usually placed either between the tenor and the vehicle or in the clause following the whole construction. It is typically expressed by verbs, verb phrases or clauses and less frequently by adjectives.

An interesting feature of these types of similes is that only masculine or neuter nouns were used as vehicles; there were no cases of feminine-vehicle nouns. On the one hand, this could be a coincidence, but on the other hand this fact gives the grounds to assume that the gender of the vehicle might have influenced the choice of the comparison marker. Though for the time being such an assumption can only be a subject of speculation until further investigations of Old English similes have been conducted.

The chronological aspect can hardly be considered an exact moment of the "birth" or "death" of certain structural types, though it shows the tendency to be widely and constantly or, for some reasons, sporadically used. The chronological appearance of the constructions in question is presented in Table 2. Since the exact date of creation is unknown and there is still much speculation concerning many of the manuscripts, the data, the data are based on the information provided by the York-Toronto-Helsinki Parsed Corpus.

As we can see from the table, the basic Old English simile construction (N / Pron + N_{dat} +*gelicost*) is typical for the whole period: Having appeared in the 6th century, it was permanently used with different variations of salient feature expression up to the Norman conquest. Solitary instances of other structural types can be logically explained. The second structural type was first used in the 7th century "Beowulf" and then, later, in the 9th century "Andreas" for which there are two possible reasons: 1) it could be another proof of the hypothesis that the author of "Andreas" read the manuscript of "Beowulf" and might have employed the construction used in it (this subject is still a matter of speculations); 2) the possible reason for deviation from a typical syntactic pattern in these cases can be accounted for by the demands of prosody since both of them are poems. As for the single similes of the third and fourth types, since both works were translations from Latin, I can presume that the structure of the original sentence must have

influenced the translator's wording choice in the target language with Latin being more developed and, correspondingly, influential at the time.

Table 2. The chronological appearance of the Old English simile of equality containing a *gelicost* component

Structural type \ Century	6-th	7-th	8-th-	9-th	10-th	11-th	12-th[1]	Total
2.1.N / Pron + N_{dat} +*gelicost*	1	5	1	6	7	3	3	26
2.1.2. N_{dat} +*gelicost*+**clause / N_{nom}P**		1		1				2
2.2.1. N +**Adj**+ *onlicost* + N_{nom}P (N_{nom}+clause)				2				2
2.2.2. N+**V**+*gelicost*+clause					1			1
Total	1	6	1	9	8	3	3	31

In terms of explicitness of the salient feature, Old English similes in question can be either open or closed as M. C. Beardsley defined them. Of the thirty one Old English similes of equality with the superlative degree of *lice* 24 are closed while the rest 7 are open; such a big amount of closed similes clearly contributes to the elimination of complexity in the worldview of Anglo-Saxons.

3. The semantic characteristics of the Old English simile with the *gelicost* component

The semantic analysis of similes is quite a challenge judging from the existence of different approaches to their classification. Most linguists group similes according to a) the vehicles, organizing the latter into thematic groups of nature, artefacts, anthroponyms, animals etc. (Gergel 2008; Riabikina 2006; Denisova 2009; Shie 2007; Moon 2008; Tomita 2008); b) the tenors and the vehicles, organising both into thematic groups (Chernousova 2014; Kondakova 2005; Kazakov 2005; Kryvenko 2006; Chudyk 2012; Yudina 2012); c) the general meaning of simile referring to a certain sphere of extralinguistic reality, singling out such groups as appearance, action, character etc. (Talko 2009; Shevchenko 2003); d) the lexical meaning of a tenor or vehicle (Rumiantseva (2007)); e) the positive/negative connotation of the ground (Veale 2009; Hao 2009). The above-mentioned classifications were worked out for the empirical data of modern languages; to the best knowledge of the author, there is no semantic classification of the Old English similes.

[1] Though the 12th century is chronologically a Middle English period, the manuscripts in questions (Paris Psalter) are considered to have been written in the Old English according to the data of University of Leicester (http://www.le.ac.uk/english/em1060to1220/catalogue/IndexFrench.htm)

The semantic analysis of similes should definitely reflect the peculiarities of all three semantic constituents which are the tenor, the vehicle and the ground. The fourth one, a comparison marker, is not crucial in this case being more of a structural element. The most important question is which simile element should be taken as the starting point of semantic classification. How should one prioritize the tenor and the vehicle; or should it be the ground the ground around which the whole classification should be centered?? The author considers all three elements to be important, interdependent and interconnected, because the choice of the vehicle depends on the salient feature, of which it should be a typical representative. The presence of the salient feature depends on the presence of topic (if there is no topic, there can be no salient feature). Finally, the distinctiveness and clarity of the topic depends on the choice of the vehicle (the more standard and typical the vehicle-bearer of the salient feature is, the more clear and distinct is the information about the topic).

I suggest that the semantic classification of simile should be done on two levels: from a macro level perspective and a micro level perspective; with neither being better than the other, but the combination of which being able to reveal a more consistent and complex picture concerning the analysed material. The macro level perspective will consider the full panorama of similes in question from a more abstract and philosophical standpoint based on certain principles, while a micro level perspective will focus on small scale combinations and the manifestation of particular tenor-ground-vehicle models. The macro level paradigm naturally incorporates several micro-level ones, but since the amount of the analysed units is rather sparse which will inevitably lead to numerous but scarcely filled subgroups, the two subgroups will be presented separately for the sake of convenience.

3.1. Macro level paradigm of the Old English Simile with the *gelicost* component.

Since the tenor and the vehicle are of equal semantic value and are the two core elements of simile, the taxonomy of Old English similes in their lexico-semantic scope can be based on two principles: **the principle of centric quality** represented by the tenor and **the principle of morphing** represented by the vehicle. The reason to single out the principle of centric quality is the fact that while creating a simile the speaker tells about a referent denoted by the tenor by means of resorting to the image of the referent denoted by the vehicle but not vice versa: The aim is to specify, centre around <u>the tenor</u>, not the vehicle. Thus, comparing the unique colour of the bird with the one of the peacock grown with joys (10), the speaker narrates about a Phoenix, not a peacock; the author's aim is to specifically account for the exclusively particular plumage of the Phoenix, for which reason he uses the image of a peacock as nothing but the means to create an eloquent picture of the former. The same is true for any other example, for instance, when the author compares the conduct of the fourteen successors with the one of hungry lions (11),

he does not narrate about the predators. He focuses reader's attention on the extremely savage behaviour of people while the lions are just an auxiliary image helping to vividly describe the ferocious and uncivilized manner of humans. Every simile is narration about the tenor, therefore, attention is focused on the tenor, it is centred around the tenor, and the aim of the simile is the precise representation of the tenor. It allows us to speak about the principle of centric quality.

On the other hand, being the means of explanation, specification, description etc. of the tenor, the vehicle is not less semantically important than the tenor. It is the etalon of the quality that it explains, specifies, describes etc. It is the vehicle with its salient feature that reveals widely-held ideas and stereotypes of the community, that gives us the grounds to analyse similes focusing on the nature of the vehicle. The tenor is almost made equivalent to the vehicle, it is almost "turned into" the vehicle. This is why the principle of morphing comes forward. In the examples, mentioned above, (10) and (11), the Phoenix and the fourteen successors look like, remind, and are similar to the peacock and lions respectively, the authors juxtapose the two as if turning the former into the latter, making both tenors zoomorphic. To achieve the goal of the centric-morphic approach to the semantic analysis of similes, covering the whole scope of the sampled examples, the need to fill the terminology gap arises, existing due to the lack of complex universal approach to their classification. For this reason, the policy of terminology-coining has been based on the analogy with existing terms describing similes and metaphors in specialized literature, the most frequent being anthropocentric/theocentric and anthropomorphic/theomorphic. Thus, the simile (10) is zoocentric and zoomorphic falling within the scope of the similes of characteristics of appearance, while the simile (11) is anthropocentric and zoomorphic, falling within the scope of similes of behaviour. The type of the scope of similes depends on the juxtaposed elements of extralinguistic reality.

The whole bulk of the sampled similes falls within the three scopes: 1) similes juxtaposing actions/states; 2) similes juxtaposing things/people / animals/phenomena; 3) similes juxtaposing situations.

3.1.1. The Old English similes with the gelicost component juxtaposing actions / states

These similes can be organized into 5 types with the further subdivision: 1) anthropocentric similes; 2) artefactocentric similes, comprising artefactocentric-zoomorphic and artefactocentric-ecomorphic ones; 3) ecocentric-ecomorphic; 4) zoocentric-creaturemorphic; 5) theocentric similes, comprising theocentric-ecomorphic and theocentric-zoomorphic ones.

3.1.1.1. Antropocentric similes juxtaposing actions/states

Such similes are aimed at shaping a precise and clear understanding of a certain action or/and its characteristics by means of two notions juxtaposition based on the similar action they can both perform. The common salient feature in this case can be the action itself or its intensity, the extent of the action or the way it is done

so it can be either explicitly or implicitly expressed by a verb or a combination of a verb and an adverb. Among them the following subtypes can be singled out:

a) anthropocentric-ecomorphic similes compare the characteristics of the actions of some inalienable parts of a person to the actions of certain natural phenomena. They highlight the intensity of the action in which case the vehicle could be substituted with the words "very intensely" (9). In this example from the 9-th century "Andreas" the author compares the flood of blood to rain clearly stressing the intensity of the flow of blood. Another example of this kind (2) is taken from "Beowulf" where the light of Grendel's eyes is compared to the dreadful light of a flame. In both cases blood and eyes, which are inalienable parts of a person, are compared to nature elements (water and fire) both being occasionally raging, intense and uncontrollable for people cannot control their blood flow or eyes shining just like they cannot control nature. Simile used for the description of Grendel is considered to be anthropocentric in this paper since the author supports those scholars' opinions (Heaney 1999; Tolkien 2014) according to which Grendel's nature was more of a human than beast, being the ancestor of biblical Cain and the son of a woman.
b) Anthropocentric-zoomorphic similes compare people's actions with those of animals to to represent the positive and negative aspects of human nature

To describe favourable deeds the images of harmless animals are employed (12) while the images of predators are used to highlight unworthy human actions (13):

(12)
…*miht somod siþiaþ sawla mid lice,fægre gefrætwed, fugle gelicast,*
'…will journey souls together with body, handsomely adorned, to a bird most like,
in eadwelum æþelum stencum…
with noble perfumes…'

The simile taken from the 9-th century "Phoenix" describes a good Christian being able to get into paradise as easily and naturally as a bird flying while in another example taken from the 11-th century "Sermo ad Anglos" undeserving and shameful behaviour is described in a comparison of men to dogs:

(13)
… & ane cwenan gemænum ceape bicgað gemæne, & wið þa ane fylþe
'… and buying a woman in one joined purchase for common use, and practicing filth upon her
adreogað, an after anum & ælc æfter oðrum, hundum gelicost þe for fylþe ne scrifað,
one after another and each man after the other to dogs most like who have no care for filth…'

c) Anthropocentric-artefactomorphic similes juxtapose a person's soul or spirit with a wheel both spinning around to depict the movement of the

soul or spirit after death thus explaining a complicated and abstract idea with a concrete and comprehendible notion:

(14)
Swa deð monnes saul, hweole gelicost; hwærfð ymbe hy selfe.
'So dead man's soul, to a wheel most like; spins around itself.'

d) Anthropocentric-creaturemorphic simile was employed to characterise the state of a sleeping person by comparing it to death. Since death is typical not only for people but for all living-beings the example taken from the 9-th century "Solomon and Saturn" has been classified as creaturemorphic:

(15)
…sorg bið swarost byrðen, *slæp bið deaðe gelicost*.
'…sorrow is the heaviest burden, sleep is most like death'

3.1.1.2. Artefactocentric similes juxtaposing action / state / process
Such similes are aimed at shaping a precise and clear understanding of the action or a property attributed to the object by explaining of what the object is capable. Four sampled similes of the kind can be divided into:

a) artefactocentric-ecomorphic type comparing weapon to ice or wax both being capable of melting (8).
b) artefactocentric-zoomorphic type comparing a ship to a bird both gliding over the water:

(16)
Is þes bat ful scrid, færeð famigheals, fugole gelicost glideð on geofone.
'The ship, with a foamy neck, to a bird most like went over the sea'

All four examples in question were found in "Beowulf" and "Andreas". The authors resort to simile in these cases not so much to describe the objects but to highlight the attending circumstances of the actions. In (16) gliding over the sea speaks about the calm sea allowing for a smooth, quiet and continuous motion while a melting weapon in (8) testifies that there was some kind of a miracle.

3.1.1.3. Ecocentric-ecomorphic simile juxtaposing action/state/process
The only simile of this kind was employed in *Phoenix* to emphasise the transience of existence and the omnipotence of God (6).

3.1.1.4. Zoocentric-creaturemorphic simile juxtaposing action/state/process
This type of simile was found in the 11-th century "Saint Oswald" comparing a horse to a mad creature to eloquently describe the movements of a sick animal in agony:

(17)
…þa wearð his hors gesicclod, and sona þær feol, wealwigende geond
'…when his horse became sick, and immediately fell there, rolling over the
ða eorðan wodum gelicost
earth, to a mad creature most like.'

3.1.1.5. Theocentric similes juxtaposing action/state/process
They help to create a didactic atmosphere teaching people to differentiate between good and bad. They can be:

a) theocentric-ecomorphic simile as in "Christ and Satan" comparing Satan's words to poison since both can easily be flown in a stream (3).
b) theocentric-zoomorphic similes allow to reveal the nature or mood of a deity that can be either positive (employing the image of harmless animals) or negative (resorting to the image of predators):

(18)
…and se hlaford ne scrifð, ðe ðæm here waldeð, freonde ne feonde, feore
'…and the lord does not care, who the army rules over, whether friend,
ne æhtum,ac he reðigmod ræst on gehwilcne, wedehunde wuhta gelicost;…
or foe, life or possessions, but savage in mind he assaults/rushes onto all, like a mad dog…'

3.1.2. The Old English similes with the gelicost component juxtaposing things/people/animals/phenomena
Such similes can be organized into 5 types with the further semantic subdivision: 1) anthropocentric-artefactomorphing; 2) ecocentric either ecomorphic or creaturemorphing; 3) theocentric being ecomorphic or situational; 4) artefactocentric-situational; 5) zoocentric either artefactomorphing or zoomorphic. These similes usually highlight a certain property or characteristic of juxtaposed notions where the ground can be explicitly or implicitly expressed by an adjective or verb the lexical meaning of which shows some property (e. g. *to glitter*).

3.1.2.1. Antropocentric similes juxtaposing things/people/animals/phenomena
They aim at pointing out a physical or psychological feature of a human, potentially leading to some consequences, by comparing a person or his/her unalienable part to a certain artefact having a typical easily decodable feature:

(19)
Swa hit awriten is on Salomonnes cwidum ðætte se mon se ðe ne mæg
'As it is written in the proverbs of Solomon, the man who cannot hold
his tungan gehealdan sie gelicost openre byrig, ðære ðe mid nane wealle ne bið ymbworht.
his tongue is most like an unprotected city, with no surrounding walls.'

In the example taken from the 9-th century "Pastoral care" the image of an unprotected city is not much of a characteristic itself but a means to highlight the reason why a talkative person can easily be subjected to risk and danger where the simile would be the answer to this question – because he/she is unprotected and vulnerable. In other words, talkativeness may cause danger. All three sampled similes of this type can be said to function like a warning against something unpleasant.

3.1.2.2. Ecocentric similes juxtaposing things/people/animals/phenomena

These similes describe physical features of a natural phenomenon or a substance without any implicature whatsoever. They are represented by two examples:

a) Ecocentric-ecomorphic simile comparing ice to a precious stone in the 9-th century "Rune Poem":

(20)
//is // byþ oferceald, ungemetum slidor, glisnaþ glæshluttur, gimmum gelicust,…
'(Ice) is very cold, and immeasurably slippery, glitters, clear as glass, to a precious stone most like…'

b) Ecocentric-creaturemorphic simile comparing sardonyx to blood was found in the 10-th century "Lapidary" to describe its colour:

(21)
Fyfta sardonix is haten se is blode licost.
'Sardonix is hot, to blood most like.'

3.1.2.3. Theocentric similes juxtaposing things/people/animals/phenomena

Such similes serve to describe a deity or a related notion highlighting its appealing features. The authors resort to the image of a nature phenomenon to stir the readers' imagination and create positive sensations. Theocentric-ecomorphic simile compares an angel to a breeze (5).

3.1.2.4. Zoocentric similes juxtaposing things/people/animals/phenomena

Two similes of this kind, found in "Phoenix," were designed to clearly and understandably describe a mythic bird's appearance. For this purpose, both an artefact and other images of animals were employed:

a) Zoocentric-artefactomorphic simile compares the phoenix's eyes to a gem in the crown (4)
b) Zoocentric-zoomorphic simile employs the image of the closest animal to the unique bird (10).

3.1.3. The Old English similes with the gelicost component juxtaposing situations

These similes can be organized into 2 types: 1) theocentric; 2) artefactocentric. The peculiarity of similes juxtaposing situations is that both the tenor and the vehicle in this case represent a situation, not a separate participant of it. They usually exemplify extended similes where the tenor and the vehicle are expressed by complex or compound sentences comparing a current and a hypothetical situation:

3.1.3.1. Theocentric simile with the gelicost component juxtaposing situations

It employs a hypothetical situation as an explanation, the reason of the result expressed by abstract notions of mercy and concord:

(22)
…ic wille eac forþ gesecgan hwelc mildsung & hwelc geþwærnes siþþan wæs siþþan
'… I intend also to set forth what mercy and what concord there has been since the advent
<se>cristendom wæs, gelicost þæm þe monna heortan awende wurden, for þon þe þa ærran
of Christianity, most like as though the hearts of men had become changed, because the
þing agoldene wæron.
previous things had been paid for.'

3.1.3.2. Artefactocentric simile with the gelicost component juxtaposing situations

This simile marks the situation centered around a broken silver dish. The absence of silver dishes at a feast is referred to as a wonder:

(23)
Gif hit swa is swa ðu sægst, ðonne is þæt egeslicre ðonne ænig oðer broga, & is endeleas
'If it is as you say, then it is more horrible than any other terror, and it is an endless
wundor,ðæm gelicost ðe on sumes cyninges hirede sien gyldenu fatu & selfrenu forsewen,
wonder,most like golden and silver vessels being scorned in the household of a certain king,
& treowenu mon weorðige.
and wooden ones being valued.'

To emphasise the unusual characteristic of the happening the author of the 9-th century "Boethius" creates an unbelievable and unreal situation for the time, to which only a wonder can be equated, as a most suitable means to explain the impossibility of the case.

The semantic characteristics of the Old English similes with the *gelicost* component according to the centric principle of classification are shown in Table 3:

Table 3. The macro paradigm of the semantic characteristics of the Old English similes of equality with the *gelicost* component.

Scope of simile / Type of simile (centric principle)	Similes juxtaposing actions / states	Similes juxtaposing things / people / animals / phenomena	Similes juxtaposing situations	**Total**
1. Anthropocentric				
1.1. Anthropocentric-ecomorphic	3	-	-	**3**
1.2. Anthropocentric-zoomorphic	4	-	-	**4**
1.3. Anthropocentric-artefactomorphic	2	3	-	**5**
1.4. Anthropocentric-creaturemorphic	1	-	-	**1**
2. Artefactocentric				
2.1. Artefactocentric-ecomorphic	2	-	-	**2**
2.2. Artefactocentric-zoomorphic	2	-	-	**2**
2.3. Artefactocentric-situational	-	-	1	**1**
3. Ecocentric				
3.1. Ecocentric-ecomorphic	1	1	-	**2**
3.2. Ecocentric-creaturemorphic	-	1	-	**1**
4. Zoocentric				
4.1. Zoocentric-creaturemorphic	1	-	-	**1**
4.2. Zoocentric-artefactomorphic	-	1	-	**1**
4.3. Zoocentric-zoomorphic	-	1	-	**1**
5. Theocentric				
5.1. Theocentric-ecomorphic	2	1	-	**3**
5.2. Theocentric-zoomorphic	3	-	-	**3**
5.3. Theocentric-situational	-	-	1	**1**
Total	**21**	**8**	**2**	**31**

Notwithstanding the fact that a thousand-year-period separates us from Anglo Saxons the similes under consideration are not difficult to interpret, even the open ones. The context provides all the necessary information, making the message relatively easy to grasp.

Similes under analysis represent almost everything people deal with in life: nature, fauna, deities, artefacts, and life situations, except for flora. As it is seen from the statistics provided in Table 3, the most frequent similes in question are anthropocentric ones especially those focusing on the actions that might stem from the fact that Anglo-Saxon society focused on people and their deeds to a greater extent than on any other aspect of life. Nature and animals, on the other hand, are represented more often (17 out of 31 cases) in the morphic part of the similes under analysis being the etalons of some salient features rather than the focus of peoples' attention. A relatively little number of theocentric similes, however, can be misleading and should not imply that Anglo-Saxon society was mainly anthropocentric with, strangely for a medieval community, deficient theological tradition. Out of 31 sampled similes only half of them concentrated on some

material wealth. The other half concerned Christian spiritual issues, though employing non-theological images, for example in (people-sheep, man-dogs, weapon-wax) whose implicatures are purely theological. Comparing people to dogs practicing filth or a sheep clearly draws the distinction between a good Christian and a bad one. Also, through the description of the ability to melt weapons the omnipotence of God is highlighted. From this we can conclude that simile is a very powerful communicative tool capable of conveying implicit ideas not reflected in the lexical meaning of a tenor and a vehicle. The fact that the number of anthropocentric similes surpasses the number of theocentric ones can be explained by a didactic character of Anglo-Saxon texts which were more aimed at shaping a conscientious Christian rather than describing God.

3.2. Micro level paradigm of the Old English Simile with the *gelicost* component.

The initial division of similes in question according to the centric-morphic principle can be specified by further classification based on correlation of particular referents, employed to form a simile, united by a motivational feature serving as a ground of the simile. This correlation could be represented by the semantic model of simile (Figure 1) with the arrow showing the direction of similarity, on the basis of which a more concrete classification can be built.

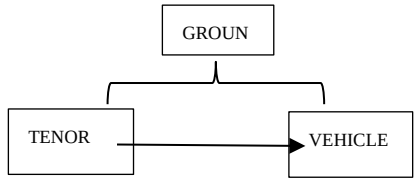

Figure 1. The semantic model of simile.

Being a commonly shared feature of the tenor and the vehicle, the ground plays a unifying role being the motive to juxtapose those two referents in the first place, therefore, the generalized scope of referents, denoted by it, might be a unifying element of the generalized scopes of referents, denoted by the tenor and the vehicle organizing them into a certain number of thematic groups. Both the tenor and the vehicle are semantically equally important, representing different aspects of speakers' viewpoints: the tenor is the phenomenon people talk about, thus showing us the priorities and subjects of their interest, while the vehicle is a phenomenon with the help of which people can eligibly specify their priorities and subjects of interest, thus, revealing to us the paradigm of images reflecting the collective consciousness. Only a particular combination of the three elements can reveal the semantics of simile, both explicit and implicit.

Whatever the aim and implicature of simile, it operates through juxtaposition of notions that, being combined to describe behaviour, functioning, appearance

etc., pass certain cultural codes representing ethnic identity reflecting certain mental and emotional features. These codes can change over time or they can stay in cultural memory: The comparison of people with sheep would not sound strange to a modern ear, but the comparison of a talkative person with a city without walls would be quite unexpected and puzzling nowadays. Thus, the classification based on centrism and morphism principle showing thematic priorities of Anglo-Saxon poetics can further be specified by the list of the motivational features, being the basis of the ground, allowing to juxtapose the tenors and the vehicles of the similes in question.

The whole paradigm of the grounds of the similes under analysis represented by concrete tenor-vehicle correlations is the following:

3.2.1. The action itself, the way it is performed, the extent to which it is performed:

3.2.1.1. The same functioning (7 out of 31):

a) light → flame with its power to shine;
b) deity → unicorn with its pure nature;
c) deity → eagle with the renovation and salvation it brings;
d) blood → rain with its heavy streams;
e) sword → ice with its ability to melt;
f) weapon → wax with its ability to melt;
g) land → wind with its ability to disappear;

3.2.1.2. the same ability (7 out of 31):

a) ship → bird (twice) with its ability to float;
b) body and soul → bird with its ability to fly;
c) dead man's soul → wheel (twice) with its property to turn around;
d) words → poison with its ability to damage;
e) deity → water with its ability to spread filling everything;

3.2.1.3. the same behaviour (5 out of 31):

a) men → dogs mating with numerous subjects;
b) people → sheep being meek and obedient;
c) people → hungry lion cubs being aggressive and merciless;
d) horse → mad creature affected by a disease, unable to control its body;
e) deity → mad dog being angry;

3.2.2. the same state (2 out of 31):

a) man → stone with its motionlessness;
b) sleep → death with its passiveness and inactivity;

3.2.3. the same physical characteristics (6 out of 31);

3.2.4. appearance:

a) bird → peacock (its look);
b) eyes → jewel with its brightness;
c) precious stone → blood (its colour);

3.2.5. property:

d) nail → steel being very hard;
e) ice → jewel being shiny;
f) angel → breeze being cool and pleasant;

3.2.6. the same psychological characteristics (1 out of 31):

a) talkative person → unprotected city being vulnerable;

3.2.7. the same characteristics of deeds (1 out of 31):

a) sin → weapon being very sharp and capable of killing;

3.2.8. the same situations (2 out of 31):

a) concord among people → change of hearts due to debt payment;
b) absence of silver dishes at a feast → preposterous kingdom;

As seen in the paradigm, it is the description of functioning and abilities of certain referents that necessitated the use of simile under consideration most frequently (14 out of 31), employing the vehicles from different scopes of reality. The behaviour, though, is always described by means of resorting to the stereotypical animals' behaviour. The thematic variety of the Old English similes in question speak in favour of the fact that Anglo-Saxons in their literary tradition concentrated on the outer manifestations of the properties of deities or people while the inner world of a person was not the focus of their attention. Most sampled similes contain concrete images as vehicles from which we can conclude that simile was not only a matter of poetics but could serve as a means to lucidly present complex ideas.

Having established the regularities of certain tenor-vehicle correlations it is possible to find out typical simile collocations in the Old English; regular combinations of vehicle with the salient feature it represents prompt us to consider the existence of some widely held ideas concerning different aspects of Anglo-Saxons' life. The material shows that two tenors (ship and floater) were used with the same vehicle (bird) on the same ground (going over the sea) as well as one and the same ground (entire melting) was used as a commonly shared feature of a weapon, ice and wax. The repetition of the vehicle as well as the ground used twice may well be considered a regularity. Thus, we can speak of an idiomatic character of the Old English phrase "to go over the sea like a bird" or we can

assume that if something disappeared an Anglo-Saxon speaker would figuratively say that "it melted like ice or wax".

4. Conclusion

Our investigation of the Old English simile of equality with the *gelicost* component has led to several conclusions: 1) structurally they represent the combination of a masculine or a neuter noun in the dative case (vehicle of a simile) with the *gelicost* component following a noun or a pronoun (tenor of a simile) with sometimes verbalized, positionally independent ground of a simile. Less frequently (in 10% of cases), when the vehicle-noun follows *gelicost*, the nominative case of the noun is employed, that is considered to be the loan translation from Latin; 2) the similes in question were productive throughout the whole Old English period; 3) from the macro-level perspective the majority of Old English similes with the *gelicost* component are anthropocentric and theocentric ones; from the micro-level perspective the majority of similes in question describe functioning or abilities of referents; 4) the repetitive tenor-vehicle combinations can speak of the idiomatic character of certain Old English similes.

References

Amodio, Mark. 2014. *The Anglo-Saxon Literature Handbook*. Hoboken: Wiley Blackwell.

Anderson, Orval. 1951. Once More: the OE Simile. *Philological Papers* 8. 1-12.

Baker, Peter. 2012. *Introduction to Old English.* Hoboken: Wiley Blackwell.

Beardsley, Monroe. 1981. *Aesthetics: problems in the philosophy of criticism.* Indianapolis: Hackett Publishing.

Cavill, Paul. 1999. Maxims in Old English poetry. Cambridge: D.S. Brewer.

Chernousova Olga. 2014. Funkcionirovanie obraznyx sravnenij v regionalnoj xudozhestvennoj proze i ustnoj rechi dialektonositelej. Avtoref. dis. na soiskanie nauch. stepeni kand. filolog. nauk : spec. 10.02.01 „russkij yazyk". Volgograd.

Chudyk, Dorota. 2012. Rosyjskie i polskie konstrukcje porównawcze z przymiotnikami nazywającymi barwę zieloną (semantyka i struktura). *Zeszyty naukowe uniwersytetu rzeszowskiego. Seria filologiczna: Glottodydaktyka* 4 : Zeszyt 76. 9-20.

Dawes, Gregory. 1998. The body in question: metaphor and meaning in the interpretation of Ephesians 5:21-33. Leiden: Brill.

Denisova Galina. 2009. *Kognitivnyj mexanizm sravneniya v nemeckom yazyke*. Avtoref. dis. na soiskanie nauch. stepeni doktora filolog. nauk : spec. 10.02.04 „germanskie yazyki". Samara.

Fisher, Olga. 1992. Syntax. In Norman Blake (ed), *The Cambridge History of the English Language*, vol. II: 1066-1476, 207-407. Cambridge and New York: Cambridge University Press.

Gergel, Aleksandr. 2014. *Obraznye paradigmy sravnenij, xarakterizuyushhix muzhchinu i zhenshhinu (na materiale sovremennoj angloyazychnoj zhenskoj prozy)*. Avtoref. dis. na soiskanie nauch. stepeni kand. filolog. nauk : spec. 10.02.04 „Germanskie yazyki". Ufa.

Gergel, Remus. 2008. Comparatives and Inversion in English: A (Necessarily) Diachronic Account. *U. Penn Working Papers in Linguistics.*, vol. 14.1. [Online]. 177-190. Available from:

https://repository.upenn.edu/cgi/viewcontent.cgi?article=1030&context=pwpl [Accessed: 30th Oktober 2018].

Heaney, Seamus. 1999. *Beowulf: A New Verse Translation*. London: Faber.

Kazakov Andrei. 2005. Obraznye sredstva i paradigmaticheskie ryady obrazov s nazvaniyami chastej tela v sovremennom anglijskom yazyke. Avtoref. dis. na soiskanie nauch. stepeni kand. filolog. nauk : spec. 10.02.04 „germanskie yazyki". Moskva.

Kondakova Irina. 2005. Obraznye sredstva, soderzhashhie toponimy v anglijskom yazyke : avtoref. dis. na soiskanie nauch. stepeni kand. filolog. nauk : spec. 10.02.04 „germanskie yazyki". Moskva.

Kryvenko Ganna. 2006. Zoosemizmi v anglijskij ta ukraïnskij movax: semantiko-kognitivnij i funkcionalno-pragmatichnij aspekti. Avtoref. dis. na zdobuttya nauk. stupenya kand. filolog. nauk : spec. 10.02.17 „porivnyalno-istorichne i tipologichne movoznavstvo". Kiïv.

Margolis, Joseph. 1957. Notes on the Logic of Simile, Metaphor and Analogy. *American Speech*, vol. 32, No. 3, 186-189. Durham: Duke University Press.

McLaughlin, John. 1983. *Old English Syntax.* Max Niemeyer Verlag: Tübingen.

Merritt, Carlton C. 2013. *Descriptive Syntax of the Old English Charters*. Berlin: De Gruyter.

Mitchel, Bruce. 1985. *Old English Syntax.* Oxford: Clarendon Press.

Moon, Rosamund. 2008. Conventionalized as-similes in English: a problem case. *International Journal of Corpus Linguistics*. 3-37. Amsterdam: John Benjamins.

Nevanlinna, Saara. 1993. The Structure of Middle English similes of equality. *Early English in the Computer Age: Explorations Through the Helsinki Corpus.* 139-170. Berlin and New York: Mouton de Gruyter. Available from: publication/228509486_Support_structure_for_linguistic_creativity_A_computational_analysis_of_creative_irony_in_similes. [Accessed 18th October 2018].

Riabikina Nataliia. 2006. Poetika ocenochnyx znachenij v proizvedeniyax i. a. Bunina. Avtoref. dis. na soiskanie nauch. stepeni kand. filolog. nauk : spec. 10.02.01 „russkij yazyk". Moskva.

Rumiantseva Marina. 2007. Tipologicheskie osobennosti komparativnyx konstrukcij (na materiale russkogo i nemeckogo yazykov). Avtoref. dis. na soiskanie nauch. stepeni kand. filolog. nauk : spec. 10.02.20 „sravnitelno-istoricheskoe, sopostavitelnoe i tipologicheskoe zykoznanie". Chelyabinsk.

Shevchenko Anna. 2003. Sravnenie kak komponent idiostilya pisatelya-bilingva v. nabokova (na materiale russko- i angloyazychnyx proizvedenij avtora). Avtoref. dis. na soiskanie nauch. stepeni kand. filolog. nauk : spec. 10.02.19 „teoriya yazyka". Saratov.

Shie, Jian-Shiung. 2007. The Semiotic Structure and Semantic Composition of English Frozen Similes. *Journal of Humanities and Social Sciences*, vol. 3, No. 1, 57-68.

Stodnick, Jacqueline, Trilling, Renee R. 2012. *A Handbook of Anglo-Saxon Studies*. Hoboken: Wiley Blackwell.

Talko Svitlana. 2009. Lingvoaksiologichna semantika antropocentrichnix metafor ta porivnyan v ukraïnskij i anglijskij movax. Avtoref. dis. na zdobuttya nauk. stupenya kand. filolog. nauk : spec. 10.02.17 „porivnyalno-istorichne i tipologichne movoznavstvo". Kiïv.

Tolkien, John. 2014. *Beowulf: A Translation and Commentary, together with Sellic Spell*. New York: HarperCollins Publishers.

Tomita, Saoko. 2008. Similes in Oliver Twist: Humanisation and Dehumanisation. *ERA, New Series*, vol.25. № 1&2, 25-42.

Tyler, Elizabeth. 2006. *Old English Poetics: The Aesthetics of the Familiar in Anglo-Saxon England*. York: York Medieval Press.

Veale, Tony, Hao, Yanfen. 2009. Support Structures for Linguistic Creativity: A Computational Analysis of Creative Irony in Similes. [Online] Available from https://www.researchgate.net/

Walker, Marilyn A. 2016. Automatically Inferring Implicit Properties in Similes. *Proceedings of NAACL-HLT*. [Online]. 1223–1232. Available from: http://www.aclweb.org/anthology/. [Accessed: 30th October 2018].

Yudina Yuliia. 2012. Modeli obrazov sravnenij v yazyke russkoj xudozhestvennoj literatury konca xx - nachala xxi vv. Avtoref. dis. na soiskanie nauch. stepeni kand. filolog. nauk : spec. 10.02.01 „russkij yazyk". Volgograd.

Dictionaries and electronic databases

DOEC = Dictionary of Old English Corpus. [1981] 2009. Antonette diPaolo Healey with John Price Wilkin and Xin Xiang (compilers), The Dictionary of Old English Corpus Web Corpus, Tei-P5 Conformant Version. Toronto: University of Toronto Center for Medieval Studies DOE Project

YCOE = The York-Toronto-Helsinki Parsed Corpus of Old English Prose. 2003. Ann Taylor, Anthony Warner, Susan Pintzuk, Frank Beths (compilers). Department of Language and Linguistic. Science University of York Heslington. York England YO10 5DD.

THE INTERPLAY OF CROSS-LINGUISTIC DIFFERENCES AND CONTEXT IN L2 IDIOM COMPREHENSION

FERRAN SUÑER
Université catholique de Louvain, Belgium
ferran.suner@uclouvain.be

Abstract
The present study investigates to what extent the effect of cross-linguistic differences on L2 idiom comprehension is modulated by the presence of a context. Sixty students of German as a foreign language (L1 French) completed a comprehension test consisting of metaphorical idioms in the L2 that differed from their L1 equivalents conceptually and formally and were presented with or without context. The results show that an increasing degree of conceptual and formal distance as well as the absence of context are generally associated with lower performance in the idiom comprehension test. However, the analysis of interactions shows that the presence of the context was especially supportive for conceptually different items, whereas the facilitative effect of formal similarity considerably diminished with increasing conceptual distance.

Keywords: idioms, metaphoric competence, cross-linguistic differences

1. Introduction

Since Lakoff and Johnson (1980) describe metaphors as mappings of conceptual content from a source domain onto a target domain that are essential for the everyday reasoning and conceptualization of the world, research into the role of metaphors has been gaining momentum in the context of foreign language learning. As a matter of fact, a significant and growing body of research exists on the acquisition of metaphoric competence, referred to as the ability to deal with metaphors and encode conceptual information in a culturally adequate way (Littlemore et al., 2014; Danesi, 2008). This overarching competence has been found to have a positive impact on different aspects of language acquisition such as sociolinguistic, textual and illocutionary aspects, as well as vocabulary acquisition (Azuma, 2009; Littlemore and Low, 2006; Cameron and Low, 1999). Acknowledging the importance of metaphoric competence at different levels of L2 learning and the necessity of leveraging the treatments that address this competence in the classroom, the question arises as to how far the L1 cultural and linguistic background of learners influences L2 metaphor comprehension and thus the effectiveness of such treatments (Hoang, 2014). Although this question has been addressed in many studies (De Cock and Suñer, 2018; Türker, 2016; Chen and Lai, 2013; Yeganehjoo et al., 2012; Ferreira, 2008; Charteris-Black, 2002; Liontas, 2002; Irujo, 1986), the interaction effects between cross-linguistic

aspects and other key factors such as the context and frequency of use on L2 metaphor comprehension are still poorly understood. Therefore, further research needs to include more complex study designs in order to gain deeper insights into the interplay of those factors and thus provide a complete picture of the variability in L2 metaphor comprehension. Consequently, the present study aims to investigate how different degrees of conceptual and formal distance between metaphorical idioms in the L2 and their L1 equivalents interact with the presence or absence of contextual information. The findings of the present study may shed light on how metaphorical expressions can be effectively introduced to L2 learners and which cross-linguistic differences need to be taken into account (Hoang, 2014).

First, we will discuss the models of L2 idiom comprehension and give an overview of studies on factors affecting L2 idiom comprehension (section 2) as well as discuss in more depth the few studies that look at the variables under investigation in the present study (section 3). We will then present a study that tested the interaction effects of cross-linguistic differences and the presence or absence of a context on L2 idiom comprehension among French speaking learners of German (sections 4 and 5). Finally, we will discuss the findings against the backdrop of previous literature, present some directions for further research (section 6) as well as pedagogical implications for teaching metaphoric competence (section 7).

2. Approaches to L2 idiom comprehension

A question that has received considerable attention in the field of L2 metaphor comprehension concerns the differences between L1 speakers and L2 learners when accessing the literal and figurative meaning of metaphorical expressions (Türker, 2016; Cieślicka, 2010). Some authors pointed out the non-compositional character of idioms and stated that with increasing L2 proficiency L2 learners are more likely to access the figurative meaning in a direct way and subsequently generally bypass the literal meaning (Gibbs, 1986). This assumption centers on the idea that proficient L2 learners tend to retrieve the meaning in a manner similar to native speakers, namely by accessing the idiomatic expressions holistically. In other words, proficient L2 learners do not necessarily need to first access literal meaning in order to infer the figurative meaning. In contrast, Cieślicka's *Literal Salience Model* (2006), which is based on Giora's (1999) graded salience hypothesis, predicts that L2 learners tend to first draw their attention to the literal meaning of the idiomatic expression and then make use of contextual clues (see also Kecskes, 2000). In this context, Giora (2003; 1999) describes salience as the degree of frequency and conventionality of a meaning that allow it to be processed first and accessed automatically from the mental lexicon, independent of contextual clues. According to this view, L2 learners favour the full compositional processing strategy to interpret the idioms rather than holistic processing. Liontas

(2002) also underscores the salience of literal meaning, describing two different phases that underlie the process of idiom comprehension in the L2: in the first phase, learners formulate hypotheses without using the context and thus rely on the linguistic mechanism. This process can be facilitated when the L1 and the L2 have similar expressions. In the second phase, the learners test the different hypotheses against the contextual information provided and progressively verify or reject them. In this vein, the *Model of Dual Idiom Representation* (Abel, 2003) postulates that L2 idiom comprehension is mainly affected by the frequency and decomposability of different idioms. Those idioms that are hardly decomposable and occur relatively frequently in the L2 input are more likely to develop a separate lexical entry, allowing direct access to their figurative meaning. The opposite is true for those idioms that have a high degree of compositionality and are rarely encountered in the L2 input, which often triggers an analysis of their components.

A major drawback of both views (compositional vs. non-compositional) is the limited empirical evidence in the field of L2 idiom comprehension when compared to L1 idiom comprehension. Only a few psycholinguistic studies on L2 idiom comprehension have been conducted to date, but they are not conclusive about the explanatory potential of both views. On the one hand, some studies (Columbus, 2010; Conklin and Schmitt, 2008) showed that L2 learners processed literal expressions faster than figurative expressions, which is in line with the patterns found in native speakers (e.g. Durrant and Doherty, 2010). These findings seem to support the holistic retrieval of idioms postulated by non-compositional approaches (Gibbs, 1994, 1986; Bobrow and Bell, 1973). However, some other authors suggest that these results could be biased because the participants were in an immersion context and had a relatively high amount of exposure to an authentic and rich L2 input (Boers and Lindstromberg, 2012). On the other hand, there is also empirical evidence for the compositional view on idiom comprehension as well as the higher salience of literal meanings. For example,., a cross-modal priming study by Cieślicka (2006) showed that idiomatic expressions prime visual targets related to their literal interpretation, which in turn accounts for the prior activation of the literal meaning of the words before accessing the figurative meaning. In the same vein, Siyanova-Chanturia and Schmitt (2011) did not find any advantage to processing idioms over novel phrases by L2 learners, which contrasted with the pattern observed with native speakers in the same experiment (see also Cieślicka et al., 2014). The authors also found that L2 learners processed idioms faster when they were used literally than when used figuratively, suggesting that the literal meaning is also accessed when inferring the figurative meaning.

Taken together, previous research supports the idea that both views complement each other in that they account for different L2 idiom comprehension conditions and explain how different factors affect this process. Whereas the findings supporting the non-compositional view suggest that increased language proficiency and the exposure to L2 in immersion contexts trigger a direct

processing of the figurative meaning of idioms (Columbus, 2010; Conklin and Schmitt, 2008; Gibbs, 1994), the other findings reported here show that L2 learners with less exposure to authentic L2 input tend to process the literal meaning of the single words first and then infer the figurative meaning (Siyanova-Chanturia and Schmitt, 2011). In this regard, the decomposability of idioms plays a major role in L2 idiom comprehension, with decomposable idioms being generally easier to comprehend (Skoufaki, 2008, Boers and Demecheleer, 2001; but see also Libben and Titone, 2008; Sprenger et al., 2006). In this sense, the so called hybrid models (Caillies and Butcher, 2007; Titone and Connine, 1999) are best suited for explaining how idioms can be situated on a continuum between non-decomposable and decomposable and how they trigger different processing strategies according to their degree of decomposability. Overall, the findings reported here suggest that, with the increase of language proficiency (or sufficient exposure to the L2), a gradual shift takes place over time in L2 idiom comprehension from a full compositional processing strategy to a non-compositional one (Cieślicka, 2015), which is modulated by many factors.

The presence or absence of a supportive context has been considered to strongly affect L2 idiom comprehension (Cooper, 1999). More specifically, the context can help learners to balance cognitive effects and cognitive efforts and reach optimal results when inferring the meaning of idioms (cf. *Relevance Theory* by Sperber and Wilson, 1995). However, some other studies downplay the importance of the context for L2 idiom comprehension and argue that L2 learners often make use of a universal pattern based on embodied experiences (e.g. using the concept of physical motion to understand expressions of time), rather than using the contextual clues (Ferreira, 2008, see also Boers et al., 2007). Moreover, another group of studies is concerned with discerning to what extent the reliance on learners' L1 linguistic knowledge and culture affects idiom processing in the L2, showing that cross-linguistic similarity, in very broad terms, facilitates L2 idiom comprehension (Chen and Lai, 2013; Charteris-Black, 2002; Liontas, 2002; Irujo, 1986). The study by Charteris-Black (2002) reveals that L2 learners manifested better performance when metaphorical expressions were conceptually and linguistically similar to their L1 equivalents. The author also stressed the importance of the frequency of exposure to L2 figurative language and intra-lingual transfer as factors affecting metaphor comprehension (see also Boers, 2003).

Further, other studies observed that L2 learners frequently use guessing from context and L1 transfer to compensate for the lack of linguistic knowledge at all levels of language proficiency (cf. Azuma and Littlemore, 2010; Azuma, 2009; Liontas, 2002), but this strategy often leads to misinterpretations (cf. Boers, 2000). Thus, learners tend to adapt their strategies depending on the difficulties they encounter and the prior knowledge available. Although the interaction effects of the different variables are essential for an increased predictability of L2 idiom comprehension models, we still know very little about them. In what follows, we

will briefly discuss the previous research on the interaction of cross-linguistic and contextual effects in L2 idiom comprehension.

3. Cross-linguistic and contextual effects in L2 idiom comprehension

Thus far, previous studies have provided valuable insights into the different factors affecting idiom processing in the L2. Although they represent an important step forwards, they neglected to examine how the different factors interact with one another. To the best of our knowledge, only three studies have been conducted to date that are concerned with the interaction of cross-linguistic and contextual effects in L2 metaphor comprehension (De Cock and Suñer, 2018; Türker, 2016; Liontas, 2002).

The study by Liontas (2002) investigated to what extent the context and type of idiom affected L2 idiom comprehension and strategy use by different groups of L2 learners (Spanish, French and German). The items were divided into three different categories (*identical*, *similar* and *different*) depending on their syntactical, semantic and conceptual similarity to the L1 equivalents and were presented with and without context. The results revealed that the presentation of a context raised the rate of correct answers in all three idiom categories and all three foreign languages tested in the study, with the identical idioms being the easiest to comprehend and the different one more difficult, respectively. The findings thus support the idea that context has an overall facilitative effect on L2 idiom comprehension. However, the study suffered from a significant flaw in that no proper distinction was made between the lexical and the conceptual levels of similarity, which could have provided more fine-grained insight into how the cross-linguistic effects came about.

The study by De Cock and Suñer (2018) investigated to what extent processing metaphorical taboo expressions is impaired when the target expression (L2=Spanish) and its equivalent in the L1 (French) do not share the same conceptual and sociocultural content. The authors made the distinction between conceptual and sociocultural aspects on the basis of the different knowledge sources needed to infer the meaning of the taboo expressions extracted from the MadSex Corpus (Pizarro, 2013). For example, the authors assumed that for expressions such as *los huevos* (‘the eggs’ = ‘the balls, the testicles’) L2 learners can guess the meaning by using conceptual knowledge arising from their embodied experiences. However, when coping with taboo expressions such as *los bebés vienen de Paris* (‘Newborn babies come from Paris’ = ‘babies are brought by the stork’; ‘sexual reproduction’), learners are expected to use very specific sociocultural knowledge that they often lack and thus are more likely to have difficulty processing metaphors. The study also looked at the availability of context as an additional factor and controlled for aspects regarding the linguistic biography of the participants (L2 proficiency level, stays abroad, etc.). The results showed that the conceptual and sociocultural distance between metaphorical taboo

expressions in the L1 and L2 do not directly affect comprehension. As to the role of context, the analysis showed that it only helps when interpreting metaphorical taboo expressions that are socioculturally different than the L1 equivalents. On the other hand, context was found to impair the comprehension of conceptually different taboo expressions. Accordingly, learners seem to use context as a source of knowledge only when they cannot use their more general conceptual knowledge. Furthermore, the authors found that some mistakes made by L2 learners when dealing with conceptually different items might be attributed to formal similarities between the target expression and other non-equivalent expressions in the L1 (e.g. the use of a reflexive verb). This suggests that formal similarity might also play a role in metaphor comprehension and should be taken into account in future research. Finally, the authors also pointed out that stays abroad and consequently better access to authentic language can facilitate processing metaphorical taboo expressions.

The study by Türker (2016) also looked at the influence of conceptual and linguistic differences, as well as the role of the contextual information in metaphor comprehension in the L2. In contrast to the study by De Cock and Suñer (2018), the author took into consideration the differences regarding both the linguistic (formal) realization of the metaphorical expression and its underlying conceptual metaphor, since two different languages sharing the same metaphor may realize it by means of different mappings (p. 31). For example, the conceptual metaphor HAPPINESS IS (DESIRED) HIDDEN OBJECT is present both in English and Korean, but the metaphorical expression *to hug/hold happiness* can only be found in Korean, even if it arises from the same conceptual metaphor as in English. The study also investigated the role of context by distinguishing different degrees of contextual information (no context – limited context – rich context) as well the influence of the frequency of the L1 equivalent expressions. The results showed that learners performed better when the L1 and the L2 were conceptually and linguistically similar. Interestingly, the descriptive statistics reveal that metaphor comprehension was generally much more affected by the linguistic distance between the L1 and L2 than by those that differed at a conceptual level. Furthermore, the context was found to facilitate comprehension only when figurative language in the L2 did not share any similarity with the L1, both at conceptual and linguistic levels. The author concluded that, in general, the influence of context declines with increasing similarity.

Although the findings from previous studies are only partially congruent with each other, the conclusions that can be drawn for further research are as follows: First, the role of context seems to vary depending on the sources of knowledge that learners can access when processing metaphorical expressions in the L2 (De Cock and Suñer, 2018; Türker, 2016; Ferreira, 2008), which stresses the explanatory potential of this factor for describing L2 idiom comprehension. Second, the linguistic distance between target expressions in the L2 and their L1 equivalents seems to play a more important role than the conceptual distance (De Cock and Suñer, 2018; Yeganehjoo et al., 2012). Consequently, further research

should explore the influence of this factor by providing a fine-grained analysis of the related dimensions. Third, the formal similarity between the target metaphorical expression and its L1 equivalent also seems to influence metaphor comprehension in the L2, since aspects such as the type of verb (reflexive, intransitive, etc.) led learners to false interpretations. Fourth, further evidence is needed to better understand to what extent other learner variables, such as exposure to authentic communication with a wide range of registers (e.g. through stays in the target countries) facilitate coping with metaphorical expressions in the L2. Against this backdrop, it is beyond doubt that further research on L2 idiom comprehension requires more complex study designs that allow an examination of the interaction between different factors. Consequently, the present study seeks to provide deeper insights into how the different degrees of linguistic similarity (e.g. conceptually identical, similar or different and formally similar or different) and the presence or absence of context interact with one another.

4. The study

4.1 Research questions

The present study aims to answer the following research questions:

1. To what extent does the conceptual and formal distance between L2 idiomatic expressions and their L1 equivalents affect comprehension?
2. To what extent does contextual information facilitate the comprehension of idiomatic expressions?
3. What is the relationship between different degrees of conceptual distance, formal distance and the presence or absence of context?

4.2 Participants

Sixty learners of German as a foreign language with French as L1 participated in the study (34 female, 26 male). They were recruited from the Catholic University of Louvain and the University of Mons, both in the French-speaking region of Belgium. The participants were students enrolled in the MA programs on modern languages or translation & interpreting. According to the internal tests and the participants' self-assessment, their proficiency level ranged from B2 to C1 according to the Common European Framework of Reference (CEFR). Over the course of the academic year, participants received an average of 30 hours of language instruction in addition to specialized courses (linguistics, literature, translation) that were also taught in the target language. With regard to the amount of exposure to authentic L2 input, 63.3% of the students (N=38) reported that they had spent at least three months in a German-speaking country, while 26.7% of the

students (*N*=16) did not have any experience living abroad (*N*=6 missing values, 10%).

4.3 Instruments

In order to investigate the effects of conceptual similarity, formal similarity and contextual information on metaphor comprehension in the L2, we created an idiom comprehension test consisting of 24 idiomatic expressions that were extracted from DWDS corpus (*Digitales Wörterbuch der deutschen Sprache*) as well as from authentic journalistic texts. The items were chosen according to the different dimensions represented in Table 1, so that every item could be assigned to the respective cell. The number of items was balanced across all categories: four idiomatic expressions in each of the three categories of *conceptual distance* (*identical*, *similar*, and *different*) for each of the categories of *formal distance* (*similar* and *different*). In order to account for the mode of presentation as a further variable (*with* and *without context*), we created two different versions of the test with the same 24 idiomatic expressions. In both test versions, 12 idiomatic expressions were presented with context and 12 idiomatic expressions without context. Those idiomatic expressions that were presented with context in the first version of the test were presented without context in the second version and vice versa. This allowed us to include the mode of presentation as a further variable without any decrease of the overall number of items for each category. Furthermore, the order of the items was fully randomized in both test versions with the aim of controlling for sources of potential measurement errors.

Table 1. Study design

	Formal difference	**Mode of presentation**	
		With context	Without context
conceptually identical (same concept)	similar	*C1SC*	*C1SN*
	different	*C1DC*	*C1DN*
conceptually similar (different concept, same domain)	similar	*C2SC*	*C2SN*
	different	*C2DC*	*C2DN*
conceptually different (different domain)	similar	*C3SC*	*C3SN*
	different	*C3DC*	*C3DN*

We distinguished three different degrees of conceptual distance between the target expressions and their L1 equivalents on the basis of Langacker's (1991: 547) definition of conceptual domain as "[a]ny coherent area of conceptualization relative to which semantic structures can be characterized (including any kind of experience, concept or knowledge system)". Accordingly, a first group of items consisted of those target expressions that used exactly the same concepts as their

L1 equivalents and were, thus conceptually identical. In a second group, we included idiomatic expressions that used concepts different from those contained in the L1 equivalents but were considered to share the same conceptual domain (e.g. body parts). This category was called conceptually similar. The third group contained idiomatic expressions that shared neither the concept nor the conceptual domain with their equivalents (conceptually different) and were assumed to be the most difficult to comprehend.

Furthermore, we made a distinction between target expressions that were formally similar to their L1 equivalents in their linguistic realization and those that were different. For example, the idiomatic expression *mit einem Fuß im Grabe stehen* ('to stand with one foot in the grave'*) and its equivalent in French *avoir un pied dans la tombe* ('to have one foot in the grave') use the same concepts, but they differ in that they are realized in the form of a transitive and an intransitive construction, respectively. In contrast, the idiomatic expression *auf die Tube drücken* ('to press the tube'* = 'to step on the gas') and the French equivalent *appuyer sur le champignon* ('to press on the champignon'*) are both realized by means of reflexive verbs and are thus formally (morphosyntactically) similar, whereas they use different concepts. Finally, all the items were presented with and without any context. Given the fact that the items with context were extracted from corpora and as such represent real usage events, the test met the claims made in previous research that metaphors should be presented in authentic language contexts rather than explanatory contexts (Boers, 2003).

Once the test was constructed, an independent rater was asked to judge if the assignment of the items to the different dimensions tested in the study was appropriate and aligned with the theory. This was the case for 95.6 % of the items. Disagreements were resolved by discussion until consensus was reached.

The participants were asked to indicate the meaning of the 24 items presented in the test and/or look for an equivalent idiomatic expression in their L1. For example, for the target expression *unter die Arme greifen* ('to give somebody a hand'), we accepted both the equivalent idiomatic expression in the L1 French (Fr. *donner un coup de main*) and the description of the concrete meaning (Fr. *aider, soutenir*='to help'). The following example with a sample solution was included at the beginning of the test in order to explain the task to the participants:

Que signifient les mots soulignés en français ? Connaissez-vous des expressions équivalentes en français ?
Par ex. : Den Nagel auf den Kopf treffen
=> Signifié : *dire vrai, deviner, faire mouche, mettre dans le mille*

What do the underlined words mean in French? Do you know equivalent expressions in French?
For example: To hit the nail (right) on the head
=> Meaning: *state the truth exactly, find exactly the right answer*.

In addition to the idiom comprehension test, the participants completed a questionnaire dealing with different aspects of their linguistic biography (e.g. proficiency level in German, stays abroad, knowledge of other foreign languages, etc.). According to previous research, these aspects are supposed to interact with the dimensions tested in the study and can therefore provide deeper insights into how L2 learners process metaphors.

4.4 Procedure

The participants were presented with both the test and the questionnaire and were asked to sign a consent form, which explained the purpose of the study and gave information about data collection and treatment as well guarantee of participants' anonymity and protection of personal data. In order to ensure that no external resources were used during the study, the test and the questionnaire were handed out in paper form. Although no time limit was set to complete the comprehension test and answer the questionnaire, carrying out the study took approximately 25 minutes.

4.5 Data analysis

The participants' answers were examined by two independent raters that were French native speakers with an advanced level of German. Every correct answer received 1 point, so that participants could receive up to 24 points for the entire test. For all items, we accepted several answers as correct, irrespective of whether they were equivalent expressions or a description of the meaning. The raters, who were trained in advance to ensure internal consistency of their judgments, agreed on 97% of the answers, which according to Cohen's Kappa (k=0.883, p<0.000) refers to "excellent agreement". In the case of disagreement, a discussion took place until consensus on coding for the final analysis was reached.

In order to analyse to what extent the main effects of the different categories and their interaction effects predict the likelihood that the participants answered correctly the items in the idiom comprehension test, a binary logistic regression with random effects (generalized linear mixed model) was conducted. The within-items factors were the different categories of conceptual distance (*identical*, *similar* and *different*), the formal similarity (*similar* vs. *different*) and the context (*with context* vs. *without context*). The individual responses in the idiom comprehension test were the dependent variable, which were coded by 0 (false) and 1 (correct). In order to control for within cluster similarity, we took the items of the test and the subjects as random effect parameters.

5. Results

The descriptive statistics of the idiom comprehension test (by-participants analysis) are shown in table 2 and expressed in mean scores. In what follows, we will first analyse the main effects of the variables and then go on to report on the two-way interaction effects.

Table 2. Results of the metaphor comprehension test by participant

Categories		Mean	SD	*N*	Mean (total)	*SD*
conceptual distance	*identical*	4.233	1.294	60	2.416	1.538
	similar	2.200	1.246	60	.	.
	different	0.817	0.813	60	.	.
formal distance	*similar*	4.650	1.505	60	3.625	1.537
	different	2.600	1.509	60		.
mode of presentation	*without context*	3.064	1.187	60	3.707	1.382
	with context	4.350	2.090	60	.	.

With regard to the influence of the conceptual distance between the L2 idioms and their L1 equivalents on idiom comprehension, the results show that the participants performed slightly better when the expressions were conceptually similar to their L1 equivalents. In fact, the highest total mean score was achieved for the items sharing the same concepts as their equivalents (M=4.233; SD=1.294), while the total mean score for the items only sharing the same domain (M=2.200; SD=1.246) and those using different conceptual domains (M=0.817; SD=0.813) were noticeably lower. As expected, the results of the binary logistic regression revealed a significant main effect of the variable *conceptual distance* on metaphor comprehension (F (2, 1281)=6.371, p=0.002). As to the different categories of the variable, the fixed coefficients for *conceptually similar* (β=−2.450, z=-2.019, p=0.044) and *conceptually different* (β=−5.125, z=-3.898, p<0.000) were significant, which indicates that the respective items were more likely to lead to false answers compared to the conceptually identical items.

Regarding the influence of the formal similarity between the L2 idiomatic expressions and their L1 equivalents on comprehension, we observed that the formally similar items were, in general, considerably easier to comprehend than their formally different counterparts. In fact, participants scored an average of M=4.650 (SD=1.505) for the *formally similar* items and M=2.600 (SD=1.509) for the *formally different* items. However, the binary logistic regression did not find any significant main effect of formal distance on metaphor comprehension (F (1, 1281)=2.501, p=0.114), which means that formally different items do not predict a reduced likelihood for correct answers. Despite this overall impairing effect of formal differences on idiom comprehension, a closer inspection of the individual mean scores reveals that this varies considerably across the different categories of conceptual distance (see Figure 1). When compared to the conceptually identical idiomatic expressions, the positive effect of formal similarity sharply declines for

the conceptually similar and different counterparts: For the conceptually similar expressions, the difference between the mean scores of formally similar and different items is approximately 0.2 points (formally similar *M*=0.283, *SD*=0.370; formally different *M*=0.107, *SD*=0.174), whereas for the conceptually different items there is practically no difference between the mean scores (formally similar *M*=0.043, *SD*=0.076; formally different *M*=0.051, *SD*=0.090). However, the two-way interaction effect *conceptual distance * formal distance* was not significant (F (2, 1281)=1.042, p=0.353), which means that conceptual differences are not associated with a significantly lower proportion of correct answers for formally different items when compared to their formally similar counterparts. In other words, formal similarity facilitates L2 idiom comprehension especially when the items are conceptually identical.

As far as the effect of contextual information on L2 idiom comprehension is concerned, the results consistently show that the presence of a context facilitates comprehension. In fact, the total mean score for the items *without context* (*M*=3.064, *SD*=1.187) is not as high as the total mean score for the items *with context* (*M*=4.350, *SD*=2.090), which suggests that presenting items with context generally facilitates idiom comprehension. As expected, the binary logistic regression found a significant main fixed effect of the variable *context* on idiom comprehension (F (1, 1281)=36.530, p<0.000). Furthermore, the analysis of the effect of the context across the different categories of conceptual distance showed that the context was very helpful, especially for the conceptually similar items when compared to their counterparts without context (see Figure 2), the mean score difference increasing by 0.2 points (with context *M*=0.300, *SD*=0.363; without context *M*=0.100, *SD*=0.156). Although to a lesser degree, the presence of a context was also found to facilitate comprehension of conceptually different idioms, the mean score for the items with context improving by approximately just above 0.1 points, compared to those items without context (with context *M*=0.136, *SD*=0.204; without context *M*=0.015, *SD*=0.027). In contrast, the context did not seem to affect the conceptually identical items the same way. In fact, the results show that the presence of a context leads to a mean score difference of less than 0.1 points compared to the items without any contextual information (with context *M*=0.565, *SD*=0.425; without context *M*=0.467, *SD*=0.431). As expected, the binary logistic regression reveals that the two-way interaction effect *conceptual distance * context* was significant (F (2, 1281)= 6.976, p=0.001). The analysis of the fixed coefficients shows that the presence of a context is associated with a significantly higher proportion of correct responses in the categories *conceptually similar* (β=0.959, z=2.451, p=0.014) and *conceptually different* (β=1.930, z=3.422, p=0.001). Thus, we can conclude that the context facilitates idiom comprehension, especially when the expressions are not conceptually identical. With regard to the interaction between the formal distance and the contextual information, the analysis showed that the presence of a context is not associated with a higher proportion of correct answers for *formally*

different items compared to the *formally similar* counterparts (F (1, 1281)=0.829, p=0.363).

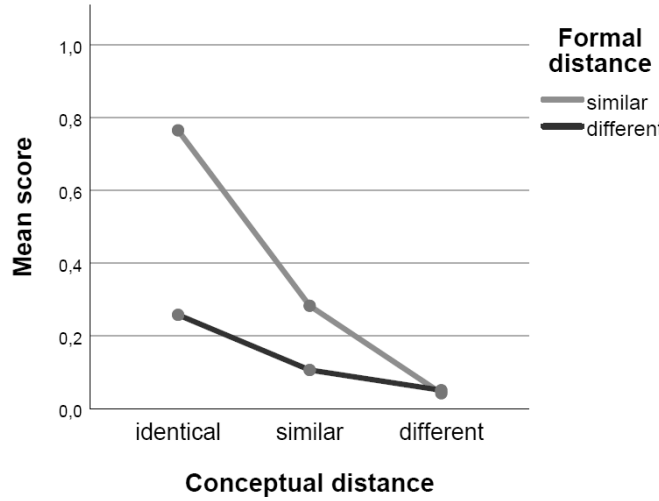

Figure 1. Results of the idiom comprehension test by categories of conceptual distance and formal similarity (by-item analysis)

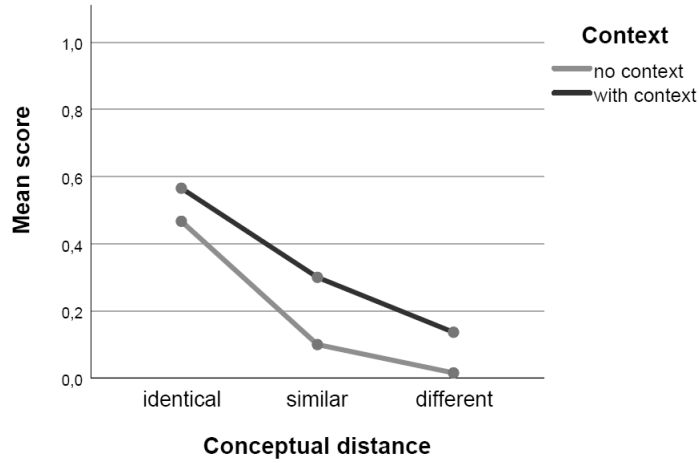

Figure 2. Results of the metaphor comprehension test by categories of conceptual distance and context (by-item analysis)

Finally, we looked at whether the performance on the idiom comprehension test can be better explained by other factors representing the participants' linguistic biography. The by-item analysis of the mean scores showed that the participants *without a stay abroad* (N=16) outperformed those *with a stay abroad* (N=38) (stay abroad, M=0.180, SD=0.245; no stay abroad M=0.188, SD=0.225). However, the

results of the binary logistic regression reveal that a stay abroad is not associated with a lower likelihood that participants answered correctly (F (1, 1281)=0.050, p=0.823). Yet, the two-way interaction effect *context * stay abroad* was not significant (F (1, 1281)=0.007, p=0.931), which means that the presence of a context does not predict correct answers by participants with a stay abroad better than for those without a stay abroad. In contrast, we found that a higher proficiency level is associated with an increased likelihood that the participants answered correctly when compared to those with a lower proficiency level (F (1, 1281)=7.924, p=0.005). In sum, the results suggest that L2 idiom comprehension is more likely to be affected by the linguistic knowledge acquired in formal learning contexts rather than by the exposure to authentic communication during a stay in the target country.

6. Discussion

We addressed three questions in our study. The first question asked to what extent the conceptual distance between L2 idioms and their L1 equivalents affects comprehension. The results showed that an increasing conceptual distance goes along with a decreasing performance in idiom comprehension. The fact that each category of conceptual distance was associated with a significantly lower proportion of correct responses in the idiom comprehension test stresses the importance of considering different levels of conceptual distance in order to gain a more fine-grained insight into its effect on L2 idiom comprehension. Interestingly, a closer examination of the wrong answers reveals that when faced with conceptually different items participants often relied on their general embodied experiences to infer literal meaning and then map it onto more abstract domains (cf. Cieślicka, 2006), even if the target expressions were conceptually different. For example, the metaphorical meaning of the idiomatic expression *Jemandem etwas in die Schuhe schieben* (‘to put something in someone’s shoes’* = ‘to lay the blame for sth. at the feet of somebody’) was often described in French as *mettre des bâtons dans les roues* (‘to put sticks in the wheels’). In general, these findings align well with previous research (Türker, 2016; Chen and Lai, 2013; Charteris-Black, 2002; Liontas, 2002; Irujo, 1986), but contrast with the results of the study by De Cock and Suñer (2018), which indicate that cross-linguistic conceptual differences do not automatically lead to an impaired comprehension of metaphorical taboo expressions. Thus, future research should focus on whether the different degrees of conceptual distance also apply to such specific topic domains. The first question also concerned the effect of formal similarity between the L2 idioms and their L1 equivalents on comprehension. The results obtained in the present study consistently show that an increasing formal difference corresponds to a decrease in idiom comprehension, which generally concurs with previous studies (Türker, 2016; Chen and Lai, 2013; Yeganehjoo et al., 2012; Charteris-Black, 2002). However, we should sound a note of caution with regard

to the consistency with previous results, since cross-linguistic formal differences have been defined in different ways: Whereas we used the term *formal distance* to refer to morphosyntactic differences between the L1 and the L2, other studies such as Türker (2016) used the term „linguistic distance“ to denominate both morphosyntactic and semantic differences in the linguistic realization of a conceptual metaphor. Thus, further work needs to be done to properly examine the effect of the different levels of formal distance on L2 idiom comprehension

Regarding the second question which concerned the facilitative effect of the context on L2 idiom comprehension, the results clearly show that the presentation of a context led to better performances. Even though these findings differ from some earlier studies (De Cock & Suñer, 2018; Ferreira, 2008; Boers, 2000), they are consistent with the studies by Cooper (1999), Liontas (2002) and Türker (2016). If we look at the answers of the participants, we can observe that participants often made use of the context, especially when similarities between the source and the target domain could not be identified (cf. Chen and Lai, 2013; Azuma, 2009; Liontas, 2002), but it seems that in some cases guessing from the context was not sufficiently effective as a strategy for L2 idiom comprehension (Boers, 2000). For example, the meaning of the expression *etwas auf den Kopf stellen* (‘To put something on the head’* = ‘to turn something upside down’) was described as *avoir quelque chose en tête* (‘to have something in mind’), which was plausible with the context about the consequences of having an in-house nursing care for an extended period, but did not match the intended meaning. At this point, the question arises as to what extent the different contexts presented in the comprehension test were equally useful or supportive to infer the meaning. In this sense, Türker (2016) demonstrated that differences in the length of the contextual information might play an important role. Thus, further work taking different aspects of the immediate context (length, informativity, etc.) into consideration will need to be undertaken.

The third question concerned the interaction effects between the different degrees of conceptual distance, formal distance and the presence or not of a context. On the one hand, we found that presenting idiomatic expressions with context was especially helpful when the expressions were not conceptually identical. Assuming that conceptual differences generally impair the metaphorical mapping, we could argue that participants made extensive use of the context, when they could not get access to the similarities between the source and target domain, which is consistent with the findings by Chen and Lai (2013). On the other hand, we found that participants benefited slightly more from formal similarity, when they were faced with conceptually identical or similar items. A possible explanation for these results may be that the formal similarity leads to a much stronger activation of the L1 equivalent and, therefore inhibits the use of other less efficient strategies such as guessing from the context or using the general embodied experiences to infer the literal meaning. Another explanation for this is that we did not control for the frequency of the L1 equivalents and this might have affected L2 idiom comprehension, being conceptually and formally items with

frequent L1 equivalents easier to comprehend that those with infrequent L1 equivalents. In this regard, some previous studies have observed that such an effect especially comes into play when idioms are presented without or with limited context (cf. Türker, 2016). Taken together, these findings support our claim that the influence of cross-linguistic differences on idiom comprehension is strongly modulated by the interactions effects between the factors.

We also looked at some other factors belonging to the linguistic biography of the participants. We found that whereas the general proficiency level was associated with better performance on the idiom comprehension test, the experience in the target country did not predict correct responses better. These findings contrast with previous research that stresses the importance of the exposure to authentic L2 input (e.g. through a stay abroad) as a factor positively influencing intra-lingual conceptual transfer and thus metaphoric competence (Siyanova-Chanturia and Schmitt, 2011; Boers and Lindstromberg, 2012; Boers, 2003; De Cock and Suñer, 2018). However, it must be mentioned that some of these studies used different methodologies (e.g. eye-tracking) and study designs (e.g. control group with native speakers), which makes a comparison of results extremely difficult. Furthermore, our study does not corroborate the findings by De Cock and Suñer (2018) who did not find any positive correlation between general proficiency level and the individual performance in the metaphor comprehension test. The reason for this rather contradictory result is still not entirely clear, but a possible explanation might be that metaphorical expressions from very specific topic domains such as taboo expressions mainly occur in informal speech, which is rarely treated in the formal L2 teaching context. Thus, the formal language proficiency level does not automatically yield a better mastery of this sort of metaphorical expressions.

Although the findings reported here add substantially to our understanding of how cross-linguistic differences and the presence or not of a context affect L2 idiom comprehension, the study has some limitations. Firstly, given that our findings are based on a limited number of idiomatic expressions (n=24), the results from such analyses could be biased by other item-related factors such as frequency or decomposability. Therefore, future work should include a higher number of items in order to balance against possible sources of unreliability. Secondly, acknowledging that translating between languages is rarely a one-way-process, it was difficult in some cases to determine which of the different L1 equivalents best matches the meaning of the respective target expressions. In fact, for some items, several options came into consideration that only partially covered the meaning of the target expression. As the choice of the L1 equivalent is essential to determine the cross-linguistic differences being tested in the experimental setting (not every potential L1 equivalent may differ conceptually and formally in the same way from the target expression), the results should be treated with considerable caution. Thirdly, it is worth mentioning that those participants that were enrolled in the master degree in translation studies might have taken advantage of their

more sophisticated translation skills in the idiom comprehension test when compared to the rest of the participants.

7. Conclusions

Overall, the present study aims to make a contribution to disentangling how the different levels of cross-linguistic differences affect L2 idiom comprehension and how they interact with the presence or absence of a context. Our findings suggest that a more differentiated treatment of idiomatic expressions, which takes into account potential difficulties for L2 learners, could help to leverage classroom interventions dealing with metaphoric competence. However, as the present study only focused on L2 idiom comprehension, further experimental investigation needs to be performed to establish whether idiom production is influenced by these factors, and thus pave the way for the elaboration of a comprehensive approach to teaching and learning idioms in the L2.

Ackowledgement

We gratefully acknowledge Harriett Jernigan for her valuable suggestions and discussions.

References

Abel, Beate. 2003. English idioms in the first language and second language lexicon: A dual representation approach. *Second Language Research* 19. 329–358.

Azuma, Masumi. 2009. Positive and negative effects of mother-tongue knowledge on the interpretation of figurative expressions. *Papers in Linguistic Science* 15. 165–192.

Bobrow, Samuel A. and Susan Bell. 1973. On catching on to idiomatic expressions. *Memory and Cognition* 1. 343–346.

Boers, Frank and Seth Lindstromberg. 2012. Experimental and intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics* 32. 83–110.

Boers, Frank, Eyckmans, June and Hélène Stengers. 2007. Presenting figurative idioms with a touch of etymology: More than mere mnemonics? *Language Teaching Research* 11 (1). 43–62.

Boers, Frank. 2003. Applied linguistics perspectives on cross-cultural variation in conceptual metaphor. *Metaphor and Symbol* 18. 231–238.

Boers, Frank and Murielle Demecheleer. 2001. Measuring the impact of cross-cultural differences on learners' comprehension of imageable idioms. *English Language Teaching Journal* 55 (3). 255–262.

Boers, Frank. 2000. Metaphor awareness and vocabulary retention. *Applied Linguistics* 21 (4). 553–571.

Caillies, Stéphanie and Kirsten Butcher. 2007. Processing of idiomatic expressions: Evidence for a new hybrid view. *Metaphor and Symbol* 22. 79–108.

Cameron, Lynn and Graham Low. 1999. Metaphor. *Language Teaching, the International Journal for Language Teachers, Educators and Researchers* 32. 77–96.

Charteris-Black, Jonathan. 2002. Second language figurative proficiency: A comparative study of Malay and English. *Applied Linguistics* 23 (1). 104–133.

Chen, Yi-chen and Huei-ling Lai. 2013. The influence of cultural universality and specificity on EFL learners' comprehension of metaphor and metonymy. *International Journal of Applied Linguistics* 23. 312–336.

Cieślicka, Anna. 2015. Idiom acquisition and processing by second/foreign language learners. In Roberto. R. Heredia and Anna Cieślicka (eds.), *Bilingual Figurative Language Processing*, 208–244. Cambridge: Cambridge University Press.

Cieślicka, Anna, Heredia, Roberto R. and Marc Olivares. 2014. The eyes have it: How language dominance, salience, and context affect eye movements during idiomatic language processing. In Mirosław Pawlak and Larisse Aronin (eds.), *Essential topics in applied linguistics and multilingualism. Studies in honor of David Singleton*, 21–42. New York: Springer.

Cieślicka, Anna. 2010. Formulaic language in L2: Storage, retrieval and production of idioms by second language learners. *Cognitive processing in second language acquisition: Inside the learner's mind*, 149–168. Amsterdam/Philadelphia: John Benjamins.

Cieślicka, Anna. 2006. Literal salience in on-line processing of idiomatic expressions by second language learners. *Second Language Research* 22. 115–144.

Columbus, Georgie. 2010. Processing MWUs: Are MWU Subtypes Psycholinguistically Real? In David Wood (ed.), *Perspectives on Formulaic Language: Acquisition and Communication*, 194–212. London/New York: Continuum.

Conklin, Kathy and Norbert Schmitt. 2008. Formulaic sequences: Are they processed more quickly than non-formulaic language by native and non-native speakers? *Applied Linguistics* 29 (1). 72–89.

Cooper, Thomas. 1999. Processing idioms by L2 Learners of English. *TESOL Quarterly* 33 (2), 233–62.

Danesi, Marcel. 2008. Conceptual errors in second-language learning. In Sabine de Knop and Teun de Rycker (eds.), *Cognitive Approaches to Pedagogical Grammar*, 231–256. Berlin & New York: Mouton de Gruyter.

De Cock, Barbara and Ferran Suñer. 2018. The influence of conceptual differences on processing taboo metaphors in the foreign language. In Andrea Pizarro (ed.), *Linguistic Taboo Revisited. Novel Insights from Cognitive Perspectives. Cognitive Linguistics Research Series (CLR)*, 379–395. Amsterdam: Mouton de Gruyter.

Durrant, Philip and Alice Doherty. 2010. Are high-frequency collocations psychologically real? Investigating the thesis of collocational priming. *Corpus Linguistics and Linguistic Theory* 6 (2). 125–155.

Ferreira, Luciane Correa. 2008. A psycholinguistic study on metaphor comprehension in a foreign language. *ReVEL* 6 (11). 1–23.

Gibbs, Raymond W. 1994. *The poetics of mind: Figurative thought, language, and understanding*. Cambrodge: Cambridge University Press.

Gibbs, Raymond W. 1986. Skating on thin ice: Literal meaning and understanding idioms in conversation. *Discourse Processes* 9. 17–30.

Giora, Rachel. 2003. *On our mind: Salience, context, and figurative language*. New York: Oxford University Press.

Giora, Rachel. 1999. On the priority of salient meanings: studies of literal and figurative language. *Pragmatics* 31. 919–929.

Grady, Joseph E. 2007. Metaphor. In Dirk Geeraerts and Hubert Cuyckens (eds.), *The Oxford Handbook of Cognitive Linguistics*, 187–213. Oxford: Oxford UP.

Hoang, Ha. 2014. Metaphor and Second Language Learning: The State of the Field. *TESL-EJ* 18 (2). Available from: http://www.tesl-ej.org/wordpress/issues/volume18/ej70/ej70a5/ [Accessed on: January 8, 2018].

Irujo, Suzanne. 1986. Don't put your leg in your mouth: Transfer in the acquisition of idioms in a second language. *TESOL Quarterly* 20. 287–304.

Kecskes, Istvan. 2000. A cognitive-pragmatic approach to situation-bound utterances. *Journal of Pragmatics* 32 (5). 605–625.

Lakoff, George and Mark Johnson. 1980. *Metaphors we live by*. Chicago: The University of Chicago Press.

Langacker, Ronald W. 1991. *Foundations of cognitive grammar*. *Descriptive application*. *Volume 2*. Stanford: Stanford University Press.

Libben, Maya R. and Debra A. Titone. 2008. The multidetermined nature of idiom processing. *Memory & Cognition* 36. 1103–1121.

Littlemore, Jeannette, Krenmayr, Tina, Turner, James and Sarah Turner. 2014. An investigation into metaphor use at different levels of second language writing. *Applied Linguistics* 32 (4). 208–429.

Littlemore, Jeannette and Graham Low. 2006. Metaphoric competence and communicative language ability. *Applied Linguistics* 27 (2). 268–294.

Liontas, John I. 2002. Context and idiom understanding in second languages. *EUROSLA Yearbook*, 2. 155–185

Pizarro Pedraza, Andrea. 2013. *Tabú y eufemismo en la ciudad de Madrid: estudio sociolíngüístico-cognitivo de los conceptos sexuales*. PhD Thesis. Universidad Complutense de Madrid.

Siyanova-Chanturia, Anna; Conklin, Kathy and Norbert Schmitt. 2011. Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research* 27. 1–22.

Skoufaki, Sophia. 2008. Investigating the source of idiom transparency intuitions. *Metaphor and Symbol* 24. 20–41.

Sperber, Dan and Deirdre Wilson. 1995. *Relevance: Communication and cognition*. Oxford: Blackwell.

Sprenger, Simone A.; Levelt, Willem J.M. and Gerard Kempen. 2006. Lexical access during the production of idiomatic phrases. *Journal of Memory and Language* 54. 161–184.

Titone, Debra A. and Cynthia M. Connine. 1999. On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics* 31. 1655–1674.

Türker, Ebru. 2016. The role of L1 conceptual and linguistic knowledge and frequency in the acquisition of L2 metaphorical expressions. *Second Language Research* 32 (1). 25–48.

Yeganehjoo, Masoomeh, Thai, Yap Ngee, Abdullah, Mardziah Hayati and Tan Bee Hoon. 2012. The influence of cross-linguistic similarities on L2 idiom production. *The Southeast Asian Journal of English Language Studies* 18. 91–107.