Sentence-in-noise perception in Monolinguals and Multilinguals: The effect of contextual meaning, and linguistic and cognitive load.

Experimenter: Charles Massingham, Durham University

Contents Page

Abstract

This study proposes a framework by which grammatically and syntactically sound sentences are classified through the perceptual measurement in noise of multilinguals and monolinguals, using an objective measure called SPERI and an interpretivist measure called SPIn, with results evaluated using Shortlist models and the BLINCS model. Hereby filling a knowledge gap on the perception of sentences that combine in varying levels of contextual meaning, linguistic load and cognitive load, this study used sentence clustering methods to find limitations of the proposed framework in determining an absolute and accurate prediction of performance between sentences in the proposed different categories, with factors such as sentence predictability and word frequency taking precedence. There were unintended findings including a relationship between the number of languages spoken and performance, proficiency in other languages decreasing performance despite being an English Native, and how mistakes by multilinguals were more semantically and phonetically influenced than monolinguals.

Introduction

The addition of background noise to the perception of speech is known to decrease performance substantially in bilinguals more than in monolinguals (see Florentine et al., 1984; Takata and Nábělek, 1990; Leather and James, 1991). So far literature on speech perception in noise comparing monolinguals and multilinguals has focused on the effect of different noise patterns, clear speech, phoneme confusion, reverberation and the characteristics of the speaker's voice (Hazan and Simpson, 2000; van Wijngaarden et al., 2002; Bradlow and Bent, 2002; Cutler et al., 2004; van Wijngaarden et al., 2004; Hedrick & Younger, 2007; Helfer, 1994; Rogers et al., 2006; Shi, 2010); the effect of context on words and sentences where the last words differentiated in predictability (Florentine, 1985; Kalikow et al., 1977; Bilger et al., 1984; Golestani et al., 2009); the effect of age of acquisition, language exposure and experience in billinguals (Bates et al., 2001; Kaushanskaya et al., 2011; Flege et al., 1997; Mayo et al., 1997; Shi, 2009, 2010; Weiss & Dempsey, 2008; Bahrick et al., 1994; Jia et al., 2002, 2006; Guion et al., 2000; Meador et al., 2000); research stating increased tone sensitivity and executive control in bilinguals (Kroll & Bialystok, 2013; Krizman et al., 2012, 2015, 2017; Bialystok, 2009, 2011) as well as phonetic identification being more difficult for multilinguals that have less linguistic experience (Krishnan et al., 2005).

There is a lack of literature that systematically combines different levels of cognitive load (sentence length), linguistic load (phonetic similarity) and contextual meaning in sentences as a framework to predict how well perceived sentences will be in monolinguals and multilinguals and what the perceptual differences are between them in noise. The previous literature has tested these factors individually, and with words rather than sentences; this study combines these factors to provide broader conclusions in the context of human communication.

This study uses the BLINCS model (Shook, 2013), and the Bayesian model Shortlist B (Norris & McQueen, 2008) to formulate hypotheses and to interpret results in the context of the proposed framework. The BLINCS model originally arose from interactive activation models that view word recognition, and ultimately speech perception, as an interactive process that involves top-down and bottom-up processing of the semantic and phonetic attributes within words and sentences (Morton,1969,; Marslen-Wilson & Welsh, 1978; Rumelhart & McClelland, 1981, 1982; Miikkulainen 1993; McClelland & Elman, 1986) that was extended, according to (Shook, 2013), to bilingual activation models (Dijkstra & van Heuven 2002; Grosjean, 1988, 1997; Li & Farkas, 2002; Zhao & Li, 2007,2010). The Shortlist Model and Shortlist B originated from bottom-up theories that viewed word recognition as a strictly bottom-up processing procedure first from the word's phonetics and selecting a word candidate from an initial search (Forster, 1976; Cutler et al., 1987; Massaro, 1989; Norris,1994; Norris et al., 1997; Scharenborg et al., 2005).

The BLINCS model describes speech perception in billinguals as an interactive process that begins with an auditory input where phonological aspects are quantified and is then processed by phonolexical (where phonetics are self-organised into a vowel-consonant structure), ortholexical (where the spelling is self-organised into a vowel-consonant structure) and semantic systems. The phonolexical and ortholexical levels share cross-language activations from both languages (Shook, 2013). On the other hand, Shortlist B describes speech perception in terms of path probabilities, where succeeding words are predicted statistically using factors such as word frequency and phoneme likelihood (Norris &McQueen, 2008).

In this study, perception of sentences was measured using an objective measure called the Sentence Perception-Error Ratio Index (SPERI), as well as with an interpretivist measure called SPIn. The BLINCS model predicts sentences with high linguistic load and no meaning to be especially difficult in multilinguals, due to the importance of semantic meaning in the interactive process, as well as interference from multiple languages on the phono-lexical and ortho-lexical levels. The length of a sentence is predicted to magnify these effects by having more words to process. Hypothesis I thus predict the removal of contextual meaning, high linguistic load and high cognitive load to individually decrease sentence comprehension and the quality of written communication in monolinguals, but more so in multilinguals. From Hypothesis I, if the quality of written communication decreases more in multilinguals, then Hypothesis II predicts more mistakes and phonetic errors as a whole for multilinguals than for monolinguals. From studies that have found language experience, immersion, exposure as well as age of acquisition to play a role in performance, Hypothesis III predicts an increase in performance from multilinguals whose native language is not English to multilinguals whose native language includes English. All hypotheses were shown to be correct.

There were 4 main unintended findings from this research. Finding I found a relationship between the number of languages spoken and performance. Finding II found that monolinguals and multilinguals categorise sentences differently to the framework proposed. Finding III found 3 main categories of perceptual difference in the sentences used between monolinguals and multilinguals. Finding IV analysed specific sentences from Finding III to observe differences in the mistakes performed by monolinguals and multilinguals.

Methods

Theoretical Framework

This paper proposes that all syntactically and grammatically correct sentences can be classified into 8 different categories that combine levels varying in contextual meaning (no meaning or with meaning), linguistic load (high or low linguistic load) and cognitive load (high or low cognitive load). This paper defines these levels for the purposes of this study alone. Sentences and level definitions used have been invented; predictability of words was done by self-judgement.

Contextual Meaning

No Meaning (NoM): The sentences make no logical sense at all. The words used in these sentences have very low predictability with each other. It is designed such it would be very difficult to guess the word from the context if not heard.

With Meaning (WiM): The sentences have logical meaning. The words used are a higher predictability than the NoM conditions.

Linguistic Load

High Linguistic Load (LinH or H): The majority of words present in these sentences individually have high functional load (Hockett, 1955), which means there exists an aspect in the word that if not pronounced well takes on a different meaning (e.g. hat, cat and sat). Adjacent words to the word of high functional load contain high functional manipulations of that word as much as grammatically or linguistically possible (e.g.hail halls healing hell).

Low Linguistic Load (LinL or L): The words in these sentences have low functional load i.e. there are very few words that sound similar to the words; adjacent words also do cannot contain deliberate functional manipulations.

Cognitive Load

High Cognitive Load (CogH or H): These sentences are 8 words long

Low Cognitive Load (CogL or L): These sentences are 4 words long

Sentence Perception-Error Ratio Index (SPERI) and Sentence Perception Indicator (SPIn)

To measure the perception of these sentences in noise, 2 measures are used:

SPERI: This index ranges from 0 to 1 (0 = completely wrong ,1 = perfectly correct with no mistakes). If a sentence scores a SPERI score of 0.5, it intuitively means that the participants correctly identified more than half of the sentence but depending on the number of mistakes made pushed the score down from 0.6 (if a participant got 60% of the words correctly identified) to 0.5. SPERI is calculated using the equation below:

$$I = \frac{W_p}{W + e - \frac{e_p}{2}}$$

Where I = SPERI score, W_p = number of correctly identified words, W= total number of words in the original sentence, e = total number of mistakes made, and e_p= number of phonetic errors made.

SPIn: This is an interpretivist binary measure of whether the sentence written is well perceived or not. This measure is intended to be a more realistic measure on whether a sentence's basic message matched the original semantically or phonetically (e.g. in the case of homophones) or both and could be comprehended (if at all possible). Sentences that phonetically matched, but not semantically, was considered well perceived (see Appendix A1).

Design

This experiment is a mixed design. Participants were sorted under three independent, between-subjects variables: ‘Linguistic Ability’ with two levels, Monolingual and Multilingual; ‘English Proficiency’ with four levels, English Native, English Native and Foreign Native, Foreign Native and English Proficient, and English Native and Foreign Proficient; and the ‘Number of Languages Spoken’. The participants were tested under 3 within-subjects independent variables each with two levels: IV_1=‘Contextual Meaning, Levels: No Meaning (Code: NoM), With Meaning (Code: WiM); IV_2=‘Linguistic Load’, Levels: Low (Code: L), High (Code: H); IV_3=‘Cognitive Load’, Levels: Low (Code: L), High (Code: H). The 3 independent variables were combined factorially together to form 8 different conditions (2x2x2). All participants did all 8 conditions.

The sentences within these conditions were measured using 4 dependent variables: SPERI, SPIn, The Total Number of Mistakes Made and The Number of Phonetic Errors.

Participants

Participants were all students (undergraduates and postgraduates) from Durham University with a mean age of 19.6 (sd=1.2). 36 females and 5 males participated in this study, totalling 41. There were 17 Monolinguals and 24 Multilinguals. Within the multilingual category, 12 were billingual, 9 were trilingual and 3 were polyglots (2 spoke 4 languages and 1 spoke 5 languages). The languages spoken in the multilingual category were German, Dutch, Mandarin, Cantonese, Hindi, Japanese, Russian, Bulgarian, Serbian, Spanish, French, Portuguese, Italian, Greek, Korean, Malay and Hungarian. All participants would have had an ILETS score of at least 6.5 (CEFR level of B2/C1, borderline high-intermediate to advanced) in English according to Durham University Entrance Requirements (The Complete University Guide, 2018). No participant had hearing problems.

Materials

6 test trial sentences and 48 experimental trail sentences (6 in each of the 8 categories) were used (See Appendix C1). The sentences were spoken by the experimenter. The sentences were then superimposed over English Human Babble. The babble used was No 19 from the SG-10 Noise-data-base developed by Dr H. Steeneken (Steeneken, 2018) which was babble in a canteen with 100 people. Sentences were counterbalanced separately for each participant in the experiment. A MATLAB program was used to present the test and experimental run. Apparatus included a computer with 2 monitor screens, 2 keyboards, 1 mouse and headphones. The experimenter had an additional laptop in front of his monitor to take measures.

Procedure

This study had received ethical approval from the Psychology Ethics Committee of Durham University and all ethical guidelines were strictly followed. After participants received an information sheet and consent form to complete, a brief introduction to the experiment was recited by the experimenter to let the participant know he/she will be completing a test trail, experimental trail and a questionnaire (see Appendix B) at the end, including debriefing. Participants wore headphones; the experimenter controlled the volume and tested the sound by playing a beep (subjects were asked if the volume was ok). Participants followed instructions including to efficiently type what they can hear as they are hearing it (to avoid serial position effects in the answers or a memory task, Murdock, 1962) and to guess when unsure. Both the participant and experimenter couldn’t see one another.

Results

The results section has been split into two parts: The hypotheses and the unintended findings. All error bars used in the graphs were standard errors customised to each condition.

Hypothesis I: The removal of contextual meaning, high linguistic load and high cognitive load will individually decrease sentence comprehension and the quality of written communication in monolinguals and multilinguals, but multilinguals will perform worse than monolinguals.

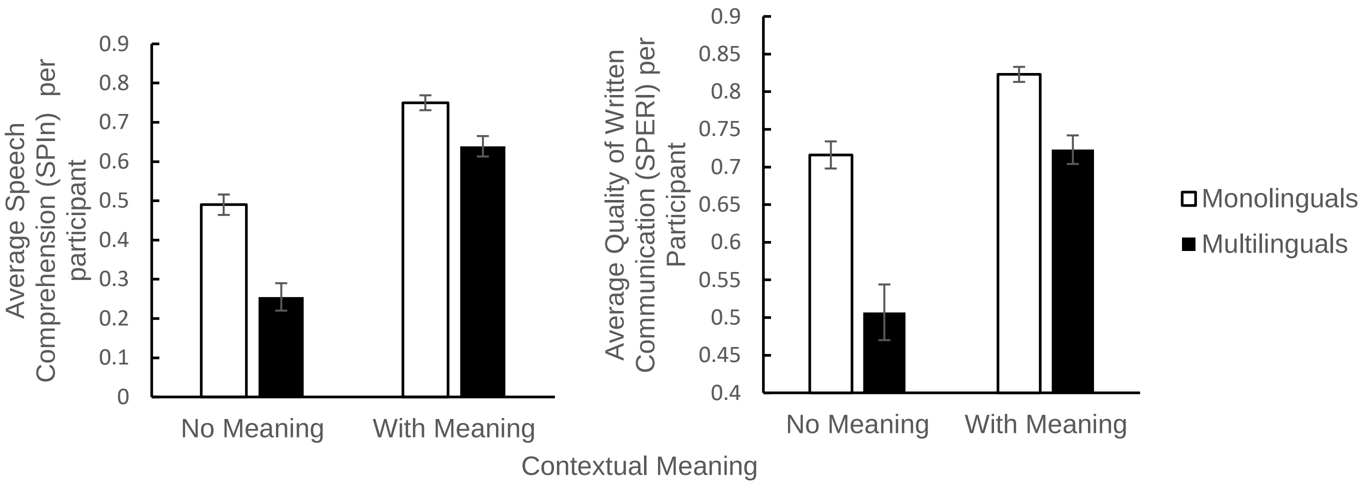

Figure 1 *The effect of contextual meaning on monolinguals and multilinguals; scores were averaged across 17 monolinguals and 24 multilinguals.*

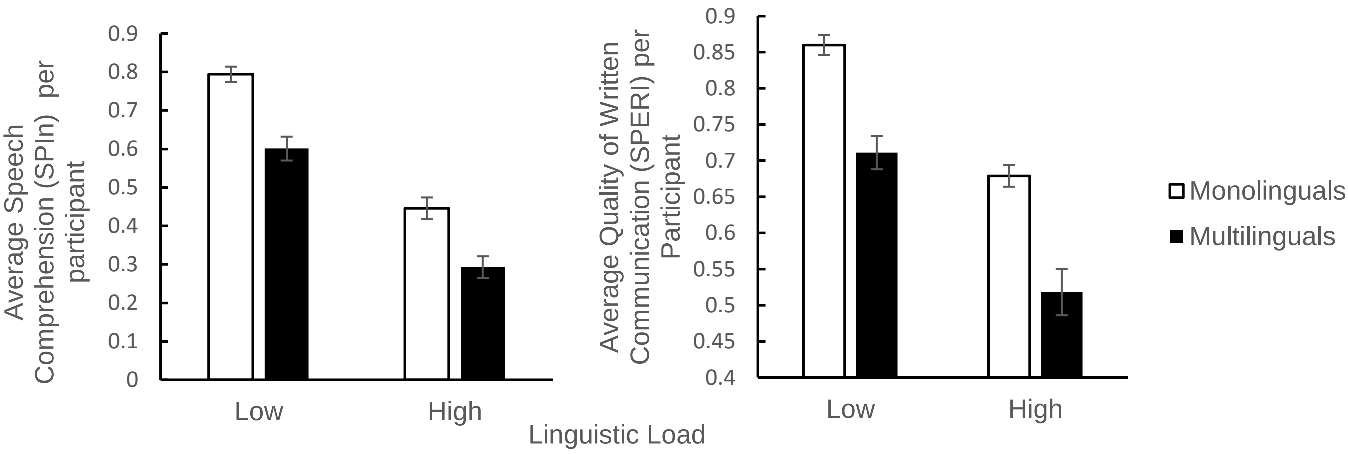

Figure 2 *The effect of linguistic load on monolinguals and multilinguals; scores were averaged across 17 monolinguals and 24 multilinguals.*

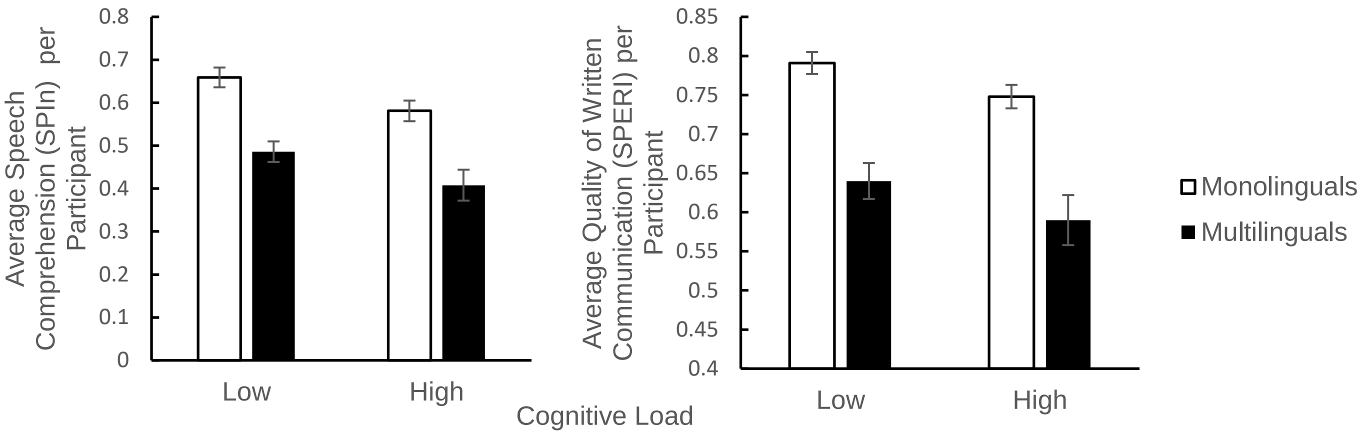

Figure 3 *The effect of cognitive load on monolinguals and multilinguals; scores were averaged across 17 monolinguals and 24 multilinguals.*

A mixed measures ANOVA was performed with Contextual Meaning (2 levels: No Meaning, With Meaning), Linguistic Load (2 levels: High, Low) and Cognitive Load (2 levels: High, Low) as within-subjects variables, and Linguistic Ability (with 2 levels: Monolingual and Multilingual) as a between-subjects variable. The ANOVA was performed separately for both the SPERI and SPIn measures.

From Figure 1, multlinguals performed worse than monolinguals in the no meaning condition in both speech comprehension and quality of written communication. The drop in performance from With Meaning to No Meaning was greater in multilinguals than in monolinguals. The effect of contextual meaning in monolinguals was confirmed significant in the ANOVA [$F(1,16)=93.536$, $p<0.001$, partial $\eta^2=0.854$(SPIn); $F(1,16)=52.9$, $p<0.001$, partial $\eta^2=0.768$(SPERI)] as well as in multilinguals, [$F(1,23)=257.234$, $p<0.001$, partial $\eta^2=0.918$ (SPIn); $F(1,23)=96.116$, $p<0.001$, partial $\eta^2=0.807$ (SPERI)]. The drop in performance was confirmed with Bonferroni-corrected post-hoc tests that found significantly greater score differences between No Meaning and With Meaning in multilinguals [0.384 (SPIn); 0.216

(SPERI), both $p<0.001$] than in monolinguals [0.26 (SPIn); 0.107 (SPERI), both $p<0.001$].

From Figure 2, Multilinguals performed worse than monolinguals in the high linguistic load condition in both speech comprehension and quality of written communication. The drop in performance from low linguistic load to high linguistic load was similar for both monolinguals and multilinguals. The effect of linguistic load in monolinguals was confirmed significant in the ANOVA [$F(1,16)=118.699$, $p<0.001$, partial $\eta^2=0.881$ (SPIn); $F(1,16)=139.364$, $p<0.001$, partial $\eta^2=0.897$ (SPERI)] , as well as in multilinguals [$F(1,23)=235.249$, $p<0.001$, partial $\eta^2=0.911$ (SPIn); $F(1,23)=365.007$, $p<0.001$, partial $\eta^2=0.941$ (SPERI)]. Similar drops in performance was confirmed with Bonferroni-corrected post-hoc tests in monolinguals [0.348 (SPIn); 0.181 (SPERI), both $p<0.001$] and multilinguals [0.307 (SPIn); 0.193 (SPERI), both $p<0.001$]

In Figure 3, Multilinguals performed worse than monolinguals in the high cognitive load condition, in both speech comprehension and quality of written communication. The drop in performance from low cognitive load to high cognitive load was small for both monolinguals and multilinguals. The effect of cognitive load in monolinguals was confirmed significant in the ANOVA [$F(1,16)=8.184$, $p=0.011$, partial $\eta^2=0.338$ (SPIn); $F(1,16)=8.736$, $p<0.01$, partial $\eta^2=0.353$ (SPERI)] as well as in multilinguals, [$F(1,23)=10.744$, $p<0.01$, partial $\eta^2=0.318$ (SPIn); $F(1,23)=13.579$, $p=0.001$, partial $\eta^2=0.371$ (SPERI)]. There was close to no difference between high and low cognitive load, which was confirmed with Bonferroni-corrected post-hoc tests in monolinguals [0.078 (SPIn), $p=0.011$; 0.043, $p<0.01$ (SPERI)] and in multilinguals [0.078, $p<0.01$ (SPIn); 0.05, $p=0.001$ (SPERI)].

One should proceed all ANOVA results in the results section with caution since for the SPIn scores, the WiM_H_L condition was found significant for the Levene´s test of equality of error variance, $F(1,39)=48.761$, $p<0.001$, and the other conditions insignificant, $F(1,39)<1.924$, $p>0.173$. For the SPERI scores, NoM_L_L, NoM_L_H, NoM_H_H and WiM_L_H were found significant, $F(1,39)<14.161$, $p<0.05$, and the other 4 conditions insignificant, $F(1,39)<2.481$, $p>0.123$.

From Figures 1,2 and 3 and their relevant ANOVA and post-hoc tests, Hypothesis I is confirmed.

Hypothesis II: Multilinguals will make more mistakes and phonetic errors than monolinguals

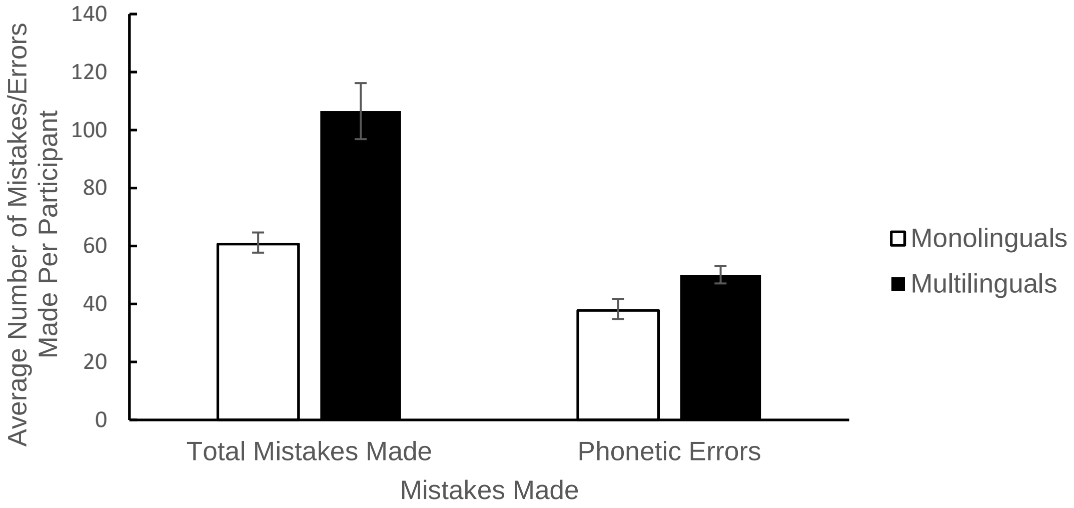

Figure 4 *The average number of mistakes (which include the phonetic errors) and phonetic errors made per participant. Scores were averaged across 17 monolinguals and 24 multilinguals.*

Figure 4 shows the total number of mistakes and phonetic errors made by multilinguals to be greater than monolinguals, with the total number of mistakes being much larger in multilinguals. To confirm these observations, an independent samples t-test was performed where the monolingual and multilingual category were treated as independent samples tested against the variables ' Total Mistakes Made' and 'Phonetic Errors'. Both the Total Mistakes Made (F=11.527, p=0.002) and the Phonetic Errors (F=6.153, p=0.18) passed the Levene's test for equality of variances. The Total Mistakes Made per participant for monolinguals (m=60.7, sd=16.3) and multilinguals (m=106.5, sd=47.4) was found to be significantly different, t(39)=3.817, p<0.001, r^2=0.272; the Phonetic Errors made per participant for monolinguals (m=37.8, sd=7.71) and multilinguals (m=50.08, sd=14.7) was also found to be significantly different, t(39)=3.139, p=0.003, r^2=0.207. As a consequence, the independent samples t-test confirms the observations and Hypothesis II.

Hypothesis III: There will be an increase in performance from Multilinguals whose native language is not English to multilinguals whose native language includes English.

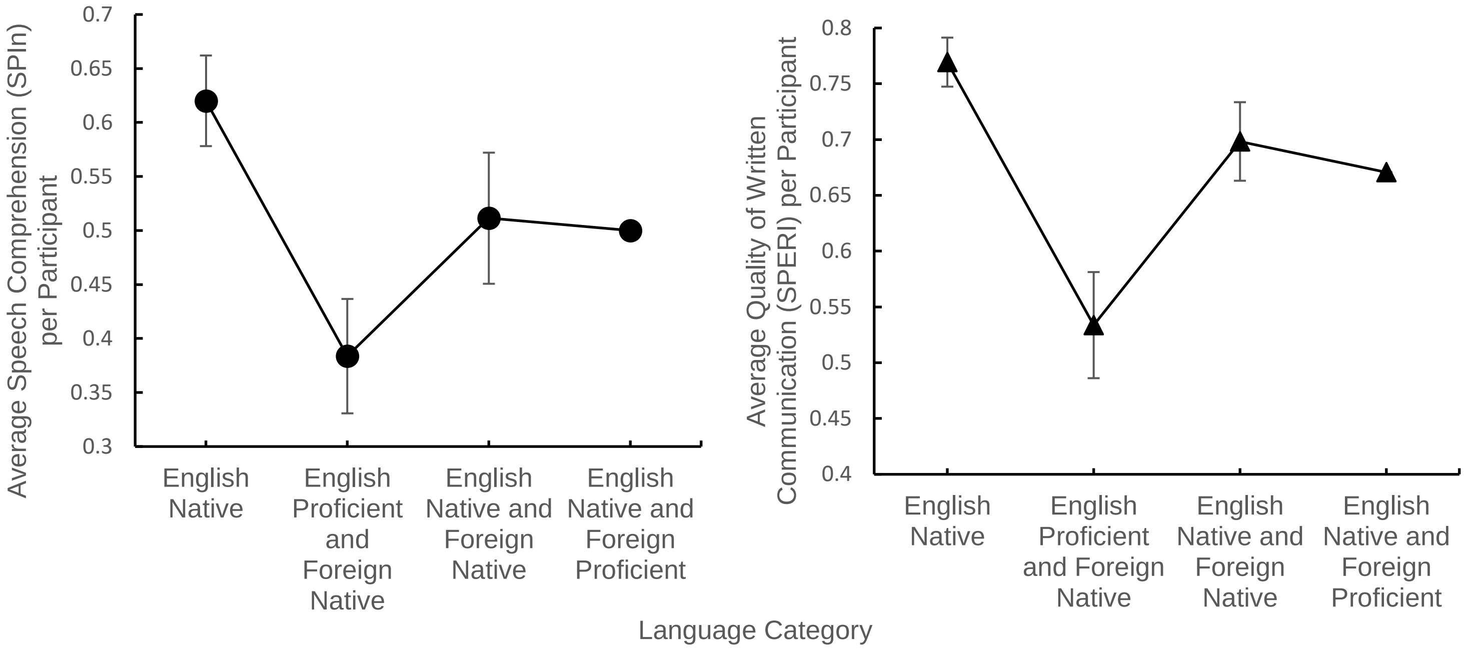

Figure 5 *The effect of Language Category on performance. Description of x-axis in Table 1. Scores were averaged across 17 monolinguals in Category 1, 11 multilinguals in Category 2 and 12 multilinguals in Category 3. There was only one multilingual in Category 4.*

Table 1: *Explanation of the Language Categories. See Appendix B for definitions of a Native Language and Proficient Language.*

Language Category	**Description**
English Native (Category 1)	Monolinguals that only speak English
English Proficient and Foreign Native (Category 2)	Multilinguals whose native languages do not include English.
English Native and Foreign Native (Category 3)	Multilinguals whose native languages do include English.
English Native and Foreign Proficient (Category 4)	Multilinguals whose native language is only English but is proficient in other languages.

Figure 5 shows Category 2 performed the worst in both speech comprehension and quality of written communication, followed by Category 3 then Category 1. Using the same within-subjects variables as in Hypothesis I, but with Language Category as a between-subjects variable, a mixed measures ANOVA was performed to confirm these observations. A clear between subjects effect was found between language category, speech comprehension and the quality of written communication [$F(3,37)=11.338$, $p<0.001$, partial $\eta^2=0.479$ (SPIn); $F(3,37)=16.797$, $p<0.001$, partial $\eta^2=0.577$ (SPERI)] . Using Bonferroni post-hoc tests, there was a significant difference between Category 1 and 2 [0.236 (SPIn); 0.236 (SPERI), both $p<0.001$], and Category 2 and 3 [0.128, $p=0.044$ (SPIn); 0.165, $p<0.001$(SPERI)]. The difference between Category 1 and 3 was not significant [$p=0.078$ (SPIn), $p=0.273$ (SPERI)]. The results from the ANOVA confirm the difference in speech comprehension and quality of written communication in Categories 1, 2 and 3, as well as Hypothesis III. Category 3 sits as an intermediate between Categories 1 and 2.

Unintended Findings

Finding I: There was a relationship between the number of languages spoken, speech comprehension and quality of written communication.

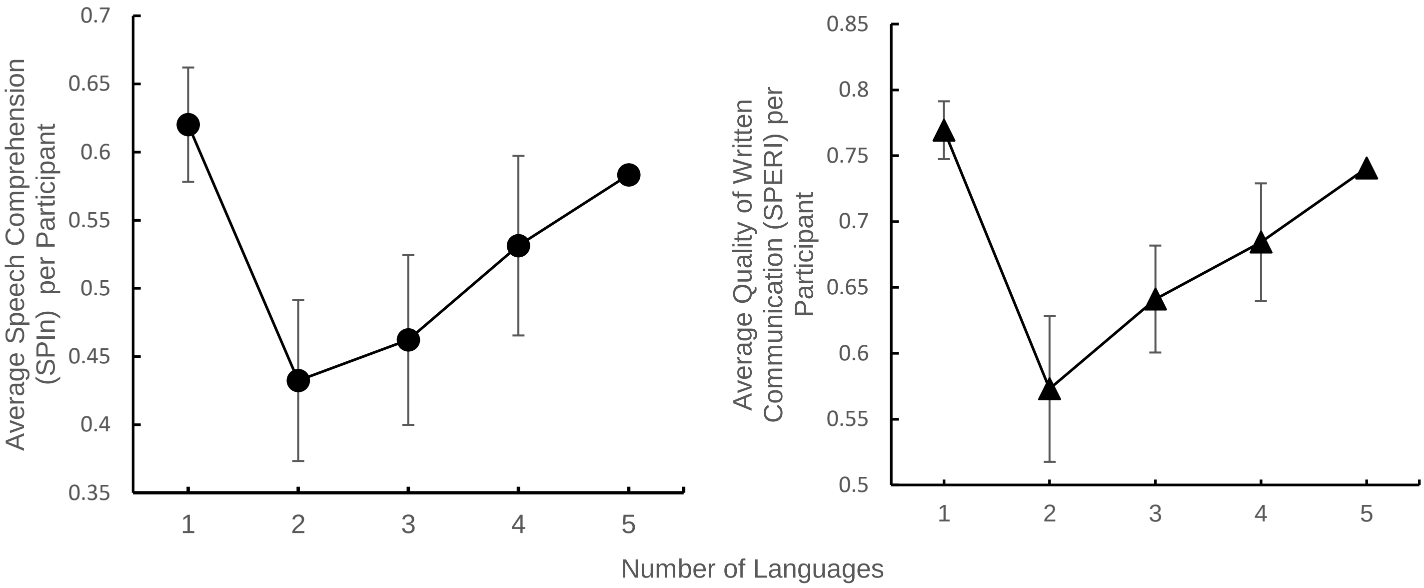

Figure 6 *The effect of the number of languages spoken and performance. All numbers include English as a language. Scores were averaged across 17 monolinguals, 11 multilinguals in Category 2, 12 multilinguals in Category 3 and 2 multilinguals in Category 4. There was only one multilingual in Category 5.*

Figure 6 shows bilinguals to perform the worst in both speech comprehension and quality of written communication, with a gradual improvement as the number of languages increase. To confirm this, the same ANOVA test as in Hypothesis III was performed, but with the number of languages as a between-subjects variable. A significant effect was found between the Number of Languages spoken, speech comprehension and quality of written communication [$F(1,36)=6.988$, $p<0.001$, partial $\eta^2=0.437$ (SPIn); $F(4,36)=6.579$, $p<0.001$, partial $\eta^2=0.422$ (SPERI)]. Bonferroni post-hoc tests found a significant difference between Monolinguals and Billinguals [0.217 (SPIn); 0.196 (SPERI), both $p<0.001$) , Monolinguals and Trillinguals [0.148, $p=0.032$ (SPIn); 0.128, $p=0.05$ (SPERI)], but no significant difference between Billinguals, Trilinguals and Polyglots ($p=1$) for both measures. Although there is suspicion of a positive monotonic improvement as the number of languages increase, there isn´t enough data to support it.

Finding II: Monolinguals and multilinguals categorised sentences differently to the theoretical framework established.

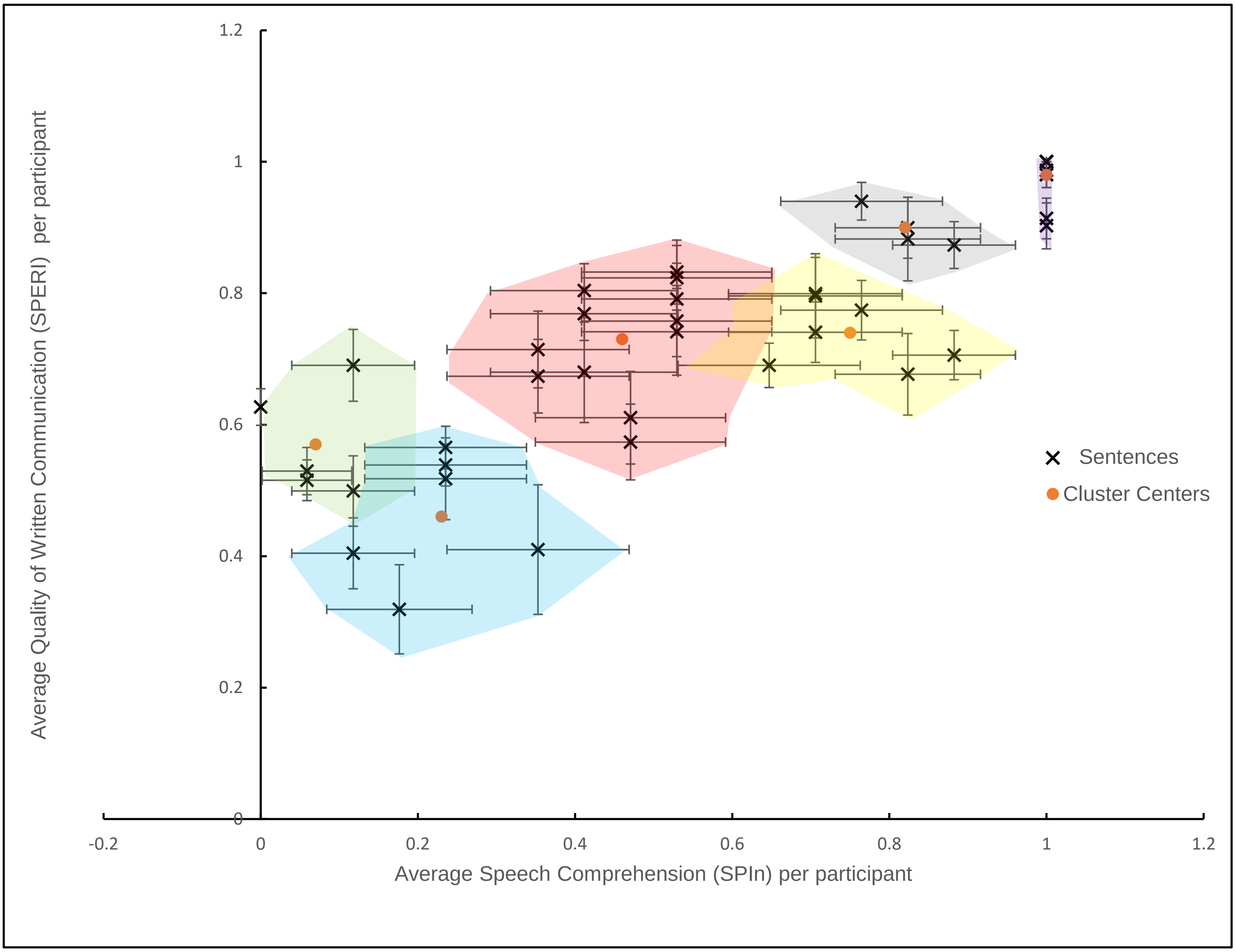

Figure 7 *All 48 sentences have been placed on a Cartesian plane with SPERI against SPIn to search for sentence clustering in monolinguals. See Appendix D1 for details on which cluster each sentence was assigned to. Scores were averaged across 17 monolinguals. Error bars represent the standard error customised for each sentence.*

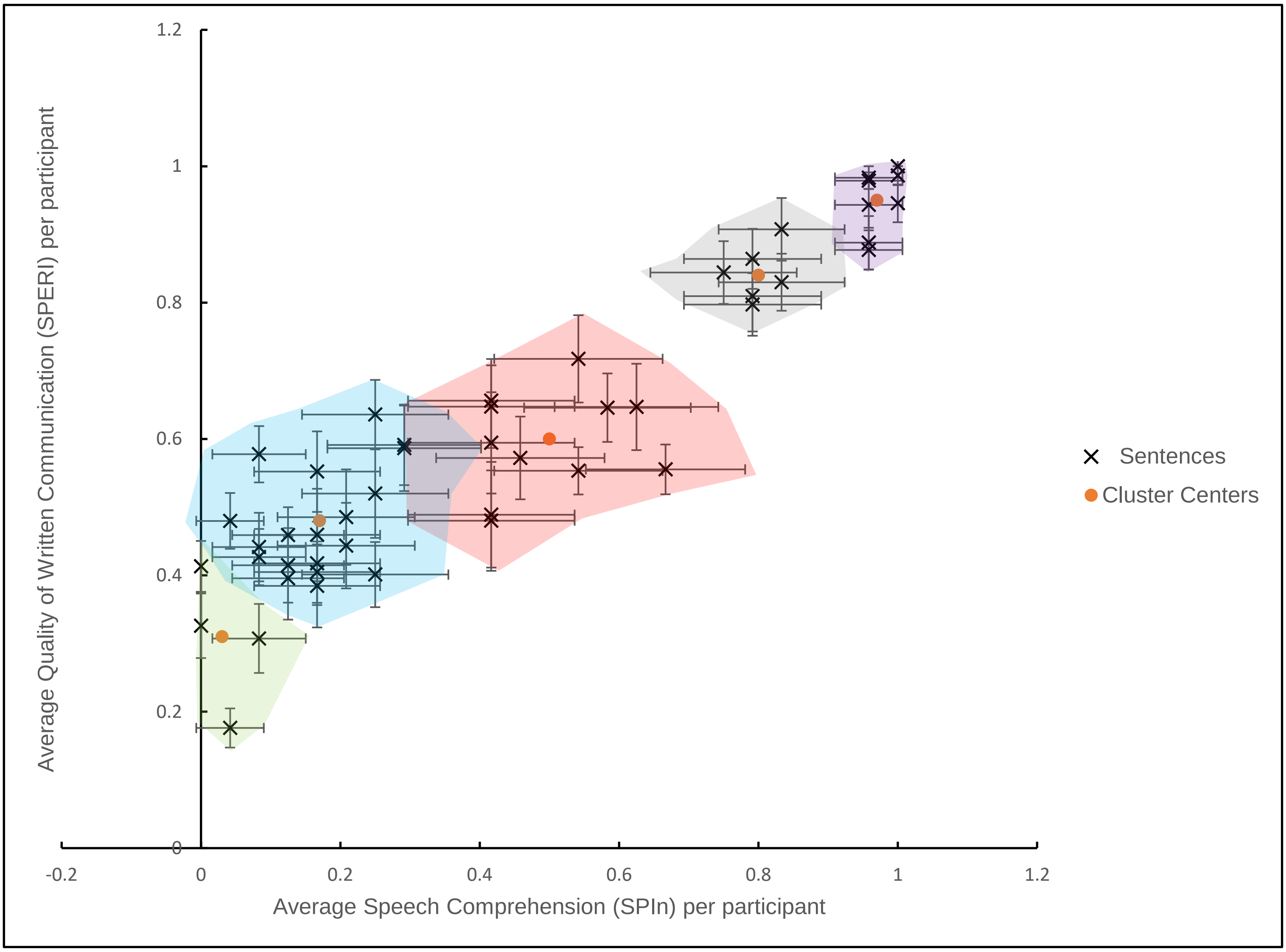

Figure 8 *All 48 sentences have been placed on a Cartesian plane with SPERI against SPIn to search for sentence clustering in multilinguals. See Appendix D1 for details on which cluster each sentence was assigned to. Scores were averaged across 24 multilinguals. Error bars represent the standard error customised for each sentence.*

Sentences have clustered into 6 different clusters in monolinguals and 5 in multlinguals which do not follow any clear pattern as laid out by the theoretical framework. To check for clusters, A K-Means Cluster Analysis was performed with ANOVA tests confirming the data's suitability for both SPERI [Monolinguals, $F(5,42)=69.806$, $p<0.001$; Fig 8, $F(4,43)=98.249$, $p<0.001$] and SPIn [Multilinguals, $F(5,42)=300.948$, $p<0.001$; Fig9, $F(4,43)=275.749$, $p<0.001$] to have data clustering. 6 Center Clusters best fitted the data in monolinguals, centring on co-ordinates (0.07,0.57),(0.23,0.46),(0.46,0.73),(0.75,0.74), (0.82,0.9), (1,0.98) respectively. 5 centre Clusters best fitted the data in multilinguals, centring on (0.03,0.31), (0.17,0.48),(0.5,0.6),(0.8,0.84), (0.97,0.95) respectively.

Clusters were coloured to show hierarchy. Clusters from worst perceived to best perceived are ordered green>blue>red>yellow>grey>purple. When comparing sentence membership to clusters in monolinguals and multilinguals, it was found that sentences have migrated from better-perceived cluster groups in monolinguals to worse-perceived cluster groups in multilinguals, resulting in the disappearance of the yellow cluster in monolinguals (see Table 2).

Table 2: *Sentence membership within each cluster and the change in membership from monolinguals to multilinguals (e.g. 14 sentences belonged to the purple cluster in monolinguals but dropped to 8 in multilinguals [red arrow pointing down], see Appendix D1 for more details). A general net migration of sentences is observed down the cluster groups from monolinguals to multilinguals.*

Cluster Group (sorted from worst perceived to best perceived)	Monolinguals	Multilinguals
green	5	4 ▼
blue	6	19 ▲
red	12	11 ▼
yellow	7	0 ▼
grey	4	6 ▲
purple	14	8 ▼

Finding III: There were 3 main categories of perceptual difference in the sentences used between monolinguals and multilinguals

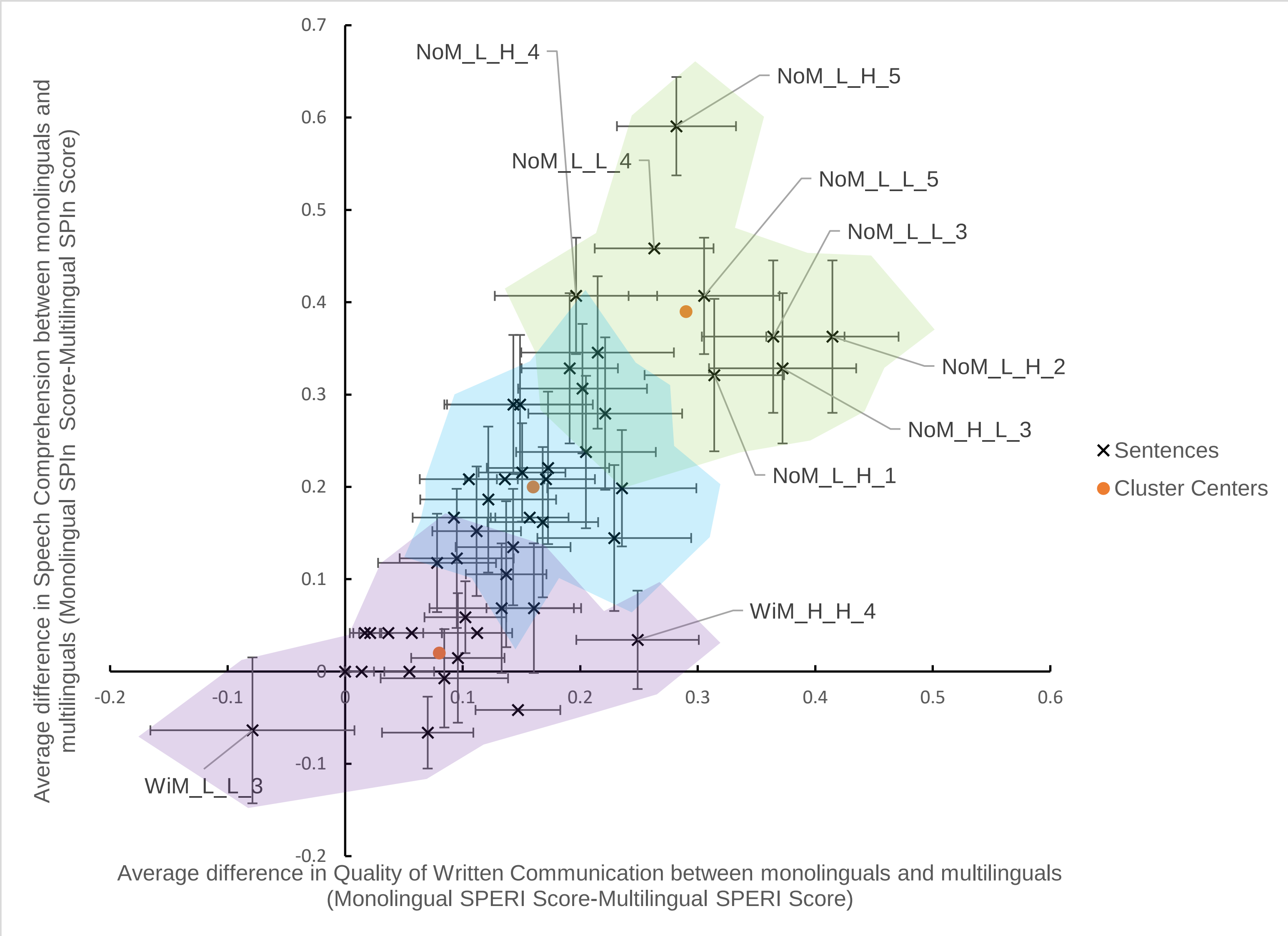

Figure 9 *Sentence co-ordinates in Figure 8 were subtracted from their matching sentence co-ordinates in Figure 7 to find net differences in scores for all sentences. 10 sentences of interest have been highlighted. 3 main clusters were found (see Table 3). Error bars represent the calculated standard error customised for each sentence using the standard deviations from Fig.7 and 8.*

Table 3: *A description of the clusters in Fig. 9. Category numbers were assigned so that colours weren't confused with the colours in Fig.7 and 8 (See Appendix D1). Figure 9 does not show how difficult sentences were, it shows the relative difference in difficulty between multilinguals and monolinguals, with the highlighted sentences being outliers in either their respective categories, or in the general trend.*

Cluster Colour	Category	Description
Purple	1	Minimal to Moderate difference in speech comprehension and quality of written communication between monolinguals and multilinguals. Multilinguals performed just as well or just as badly as monolinguals, though there are sentence outliers that have noticeable differences (highlighted in Fig. 9)
Blue	2	Moderate to Large Difference in speech comprehension and quality of written communication between monolinguals and multilinguals. Relative to the monolingual's overall performance on the sentence, multilinguals did worse.
Green	3	Large Difference between monolinguals and multilinguals. Relative to the monolingual's overall performance on the sentence, multilinguals did far worse.

A K-Means Cluster Analysis was performed to confirm the 3 clusters. ANOVA tests found the appropriateness of clustering to be significant [SPERI, $F(2,45)=37.813$, $p<0.001$; SPIn, $F(2,45)=123.743$, $p<0.001$]. 3 main cluster centres were found being (0.08,0.02), (0.16,0.2), (0.29,0.39) respectively. The classification of these clusters made it easier to find sentences that caused the biggest difference in perception between multilinguals and monolinguals for further investigation (see Finding IV).

Finding IV: Multilinguals created new words out of the phonetics of the original words, as well as new semantic content (see Table 4).

Table 4: *A descriptive analysis of sentences identified from Finding IV. Percentage shows the percentage of participants that got this sentence wrong. Numbers in brackets show repetitions (e.g. (b)= b participants wrote this). Frequency was measured using the NOW Corpus* (corpus.byu.edu, 2018), *numbers refer to the number of occurrences in a 5.9 billion word Corpus from newspapers and magazines since 2010. (as reference points; the word 'the' = 354,288,885, ' People' = 10,349,562, 'Good' =4,141,062, 'Feel' =1,397,406). All words were very low frequency (except words such as 'of', 'from', 'and', 'with', 'was', 'a', 'is', 'inside', 'here')*

Sentences	Monolingual	Multilingual
NoM_L_L_3 **Wolves distribute excessive listings** **Number of occurrences in 5.9 billion Corpus** **Wolves = 47,502** **Distribute = 56,011** **Excessive = 86,654** **Listings = 89,574** **MONOLINGUALS** **52.9%** Mistakes Overview: All made phonetic mistakes on the first word and last word. Majority of mistakes were the same mistakes, there is consistency in the type of errors made **MULTILINGUALS** **83.3%** Mistakes Overview: Phonetic mistakes made on first, second last and last word. New word creation from phonetics (e.g. Ballistics, Waltz, gold, woods, ball, halls), different word forms taken from the original (e.g. excess, access, listening). Some perceived 3 or 5 words. Most mistakes made in words that can be shortened to another word.	walls distribute excessive listings (6) walks distribute excessive listings wolves distribute excessive instincts walls distribute accesibalistics	walls distrubute excessive listings walks distribute excessive listings wolves distribute excessive instincts wolves distribute excessive blistings olds distribute excessive listings distribute excessive listings ores distribute excessive listening wolves distribute accessible instincts (2) all distributes excess wall distributes excess distinct woods distribute excess walt distribute excessive ballistics waltz distribute excess bliss

		wolf distribute access to ballistic wolf distribute access abilistics woolfs distribute instincts ball distribute accessible instincts halls distribute accessebilistics gold d accessible d
NoM_L_L_4 **Openness rewards quiet marathons** **Number of occurrences in 5.9 billion Corpus** **Openness = 25,974** **Rewards = 69,549** **Quiet = 177,104** **Marathons = 9513** **MONOLINGUALS** **0.06%** Mistake Overview: Phonetic mistake reduces word from noun to verb (e.g. opens) **MULTILINGUALS** **46.8%** Mistake Overview: Phonetic mistakes made with reward and quiet. New words created from phonetics (e.g. water, open this, fun, really). Some changed the semantic meaning to the beginning of an opinion (is quite fun, was really a). Word shortening observed with 'Openness' to ' Open'.	opens rewards queit marathons	openess rewards quite marathons (2) openess awards quiet marathons openness rewards quiet maphones openness revolts quite marathons openess remotes quiet marathons open this water quiet marathons open this reward quiete marathons open that water is quite fun openness was really a s

NoM_L_L_5 **Silver chaos enchants poems** **Number of occurrences in 5.9 billion Corpus** **Silver = 267,978** **Chaos = 106,281** **Enchants = 320** **Poems = 35,635** **MONOLINGUALS** **29.4%** Mistake Overview: Phonetic mistakes made with enchants (in chance), invention made (the filmers) **MULTILINGUALS** **54.2%** Mistake Overview: New words created from original phonetics (super, enhance, trans, silvia, transpolar, in chance, chance, pillows). Invented word (over. Majority did not perceive 4 words,but 1,2,3,5,6 or 9. Repetition of words, as well as words that have phonetic similarity (Chaos, silver and super)	silver chaos enchant poems silver chaos enchants and poems silver chaos in chance poems (2) silver chaos enchants the filmers	silver chaos enhance poems silvia chaos enchant poems silver chaoes enchanced poems silver chaoes over chant of poems silver chaos inchanced chaos enchants poems Silver chaos enchants poems super chaos in trans poems super chaos in trans poem silver chaos in transpolar so the chaos in chance poems chance pillows chant poems chance
NoM_L_H_1 **Extinction of purple corpses occurs from exquisite breath.** **Number of occurrences in 5.9 billion Corpus** **Extinction = 35,126** **of = 162,109,413 [HIGH]** **Purple = 55,927** **Corpses = 14,282** **Occurs = 70,096** **From = 28,165,408 [HIGH]** **Exquisite = 21,330** **Breath = 100,620**	extinction of purple corpses occurs from exquisite breaths extinction of purple curses occurs from exquisite breathe extinguish of purple corpses occurs from exquistite breath	extinction of purple corpses occurs from distinguished breath extinction of purple corpses occurst exquisite breath extinction of purple corpses occurs from exquisite breasts exctinction of purple corpses of excuisite breath

MONOLINGUALS 70.6% Mistake Overview: Phonetic mistakes made with first and last word (extinguish, brain), purple (herbal), corpses (curses). New words made from phonetics (pavelled horses). Repetition of phonetically related words (extinguish, purple). However, words chosen are contextually relevant and semantically feasible. **MONOLINGUALS 87.5%** Mistake Overview: Phonetic mistakes (breasts,birth,herbal New words made from phonetics (distinguished, duress, curses, horcruxes, blessed, Oscars, exhibit, press, extinguished, death, breads) Words chosen are not contextually relevant or semantically feasible. Basic structure of the sentence has been taken apart, some have been reworded to take on new semantics (e.g. extinction of herbal occurs because of death, extinction of corpses happens with exquisite breath). There is evidence of new semantic creation.	extinction of herbal corpses occurs from exquisite breath (2, one with breaths) extinction of purple corpses extinguish from exquisite breath extinct of purple corpses occurs from exquist breaths extinction of purple corpses causes breath extinction of purple corpses extinguishes purple breath extinguish of purple corpses occur from exquisit brains extintion of pavelled horses occurs excuisite breath extinction of purple exquistite breath	exstintion of purple courpses occures with duress extinction of purple corpses curses birth extinction of horrocruxes extends of exquisite breaths extinction of purple corpses purple breath extinction of purple courpses excludes fom purple breath extinction exquiste breath existicting of corpses occurs from excusit breath extension of purple corpses comes from extinction of corpses extinction of blessed extinction of herbal oscars exhibit the press extincrion of habo corpses happens with exquisite breath extinction frm purple breathe extinguished extinctinction of herbal occurs because of death exstincts of corpses of extinction of breads

NoM_L_H_2 **Tall coal bowls poll** **Number of occurrences in 5.9 billion Corpus** **Tall = 115,844** **Coal = 250,847** **Bowls = 31,031** **Poll = 231,234** **MONOLINGUALS** **47.6%** Phonetic mistakes made (paw, core, pulls, walls, hall), Repetition of phonetically similar words (pulls and paw, call and core) and one instance of word switching (ball paws) **MULTILINGUALS** **75.0%** Phonetic mistakes are more extensive, more mistakes made with plurals (paw,calls,halls, ball, pause, pawled, pour, boars,pore,tore core stoll, hall, crawl, poor). New words made from phonetics (bork, horse, thorne, goes, claws), many instances of word order errors (poll ball, balls call, pause call).	tall call ball pau tall call balls paw (2, pore used instead of paw) tall call ball core pore tall call pulls paw tall call paws ball tall call walls hall tall hall balls paul	tall call balls paw tall calls ball paul tall halls balls pall tall coll pause call tall balls call hall tall call balls pawled tall call balls pour talk ball pause paw thaw calls balls pore tall call ball pause tall coal boars pore talk poll stoll ball talk hall balls bork talk horse pall tall crawl poor balls thorne core pause ball tall holes goes tore claws bore core
NoM_L_H_4 **Agitated persons with reflected, spiky and musical surfaces** **Number of occurrences in 5.9 billion Corpus** **Agitated = 15,261** **Persons = 339,533** **With = 42,594,197 [HIGH]** **Reflected = 115,011** **Spiky = 3299** **And = 153,492,230 [HIGH]** **Musical = 233,437** **Surfaces = 36,199**	agitated persons with reflected spiky musical surfaces (2) agitated persons with reflective spiky and musical surfaces (2) agitated persons with reflected and musical surfaces agitated persons with reflectant spiky and musical surfaces	agitated persons with reflected spiky musical surfaces agitated persons with reflective spikey and musical surfaces agitated persons with spiked surfaces agitated persons with reflected agitated musical purposes

MONOLINGUALS **74.6%** Phonetic mistakes made, (reflective, reflectant,) new words made from phonetics (amusing, sparkly,music, purposes) there is some evidence of invention (intelligent, spikey and music glasses)and new semantic formation (claim from purposes) **MULTILINGUALS** **70.8%** Phonetic mistakes are more extensive (reflective, reflects spiked, person). New words made from phonetics (purposes, sarum, senses,sparky,despite circuses,affected, educated) word repetition (agitated) Words created from the phonetics of other words (shiny), sentence structure broken, some only perceived a few words. There is evidence of forming new semantics (with their music circuses, despite the circles)	agitated with persons reflected with spikey amusing surfaces aggitated persons with musical surfaces ajtated persons from musical and sparkly surfaces agitated persons with spiky musical surfaces aggitate persons with spikey and musical surfaces adjitative persons with intelligent spikey and music glasses adjitated person claim from purposes	agitated persons were affected sarum and musical senses agitated parcels which reflected sparky and musical surfaces adgitated persos reflected spiky and shiny surfaces agitated persons with spiky surfaces agitated with persons with musical surfaces educated person with their music circuses ajetated persons with spikey breath balabd musical circuis aditative agicated persons agitated person reflects purposes agited persons musical adjeted c misi musical circus agitated adjetative dispite the cirles

NoM_L_H_5 **A delivery of underwater tigers enraged spiritual incense** **Number of occurrences in 5.9 billion Corpus** **A = 126,325,848 [HIGH]** **Delivery = 355,192** **of = 162,109,413 [HIGH]** **Underwater = 38,561** **Tigers = 93,057** **Enraged = 12,938** **Spiritual = 110,579** **Incense = 4828** **MONOLINGUALS** **58.8%** Phonetic errors were made (deliver, enrage, and raged, inscents, enrages) and a new word was made from phonetics (sea, sense). There is some consistency of answers too **MULTILINGUALS** **87.5%** More phonetic errors were made (enrage, enrave,and rage , tiger, incest) new words formed from phonetics (senses, in chance,strange, insects,insults,enhance, spell, rain, sea, encends [ascends]). The word water in many cases has been taken from the original word underwater and used separately and used with other words (e.g. sea water) or has taken other forms that involve water (e.g. storm, rain). There is evidence of attempts of forming new semantic meaning (enhance the chance of storm, under water tiger ascends incense, about to water tigers) as well as a rephrasing of original semantic (a delivered tiger of rage)	a deliver of underwater tigers enraged spiritual incense a delivery of underwater tigers enrage spiritual incense (4) a delivery of underwater tigers and raged spiritual insense (2) a delivery of underwater tigers enraged spiritual inscents a delivery of underwater tigers spiritual insense a delievery of underwater tigers and enrages sea sense	a delivery of underwater tigers enrage spiritual incense (3) a delivery of underwater tigers and rage spiritual senses a delivery of underwater tigers enraves spritual incense a delivery of underwater tigers enraged spiritual insults a delivery of underwater tigers and enraged spiritual incense a delivery of underwater tiger in chance a delivery of underwater tigers and strange inssects a delivery about to water tigers spiritual incest a delivery of underwater tigers and spirital instincst a delivery under water enhance the chance of storm a delivery of underwater tigers a delivery of underwater tiger and rage spiritual incence a deliver to underwater was spell

		an underwater tigers rain spiritual and sense a delivered water tiger of rage spiritual and sense a delivery of water a delivery of underwater tigers enraged the spiritual sense a dilivery of under water tiger encends incense a delivery of underwater tiger sea water

NoM_H_L_3 **Caught cops fought caps** **Number of occurrences in 5.9 billion Corpus** **Caught = 424,867** **Cops = 95,409** **Fought = 179,957** **Caps = 104,965** **MONOLINGUALS** **64.7%** Phonetic errors were made (cot,cats,capse,traps, Thought, quart) New words made from phonetics (cart,forse,fots,force, Lot) and some phonetic repetition (quart and court) **MULTILINGUALS** **91.6%** More Phonetic errors were made (cat,cap,cats,quart,cot, Cots, thought) and more New words made from phonetics (corpse, cough, corps, fox, quote , flaps, cold, happs, sord, cod, forts, cows, cocks, cut,cups, cox, thaves, black) and some phonetic repetition (quart and caught). One should note the number of animals that have been mentioned in a sentence that had no animals (cat, fox, cows, cocks). There is some evidence of themed words that are semantically related (in the case of animals).	caught cot fought caps (3) caught cot fought cats caught, capse, thought, traps caught faught copes cat cught fots forse caps quart cops cart cats caught cops force caps quart lot court caps	caught fought cat cap quart caught fort cats cord corpse fought cats courts cops fought cats court corpse fought caps (2) court caught force cafs quote cot fuoght cats cought caugh fought cats cough corps fought cats court corpse caps quote cots through cats cort corpse fox caps quart cot fort caps cought thought cat flaps cough cought cold happs sord cod forts caps caught cot caps court called cows cut quote cocks fought cat courts cox fought blac cats court c c cups cut cat thaves

WiM_L_L_3 **Smoking here is forbidden** The strangest responses were found in this sentence. This sentence has proven to be an extreme outlier in its category. **Number of occurrences in 5.9 billion Corpus** **Smoking = 127,415** **Here = 3,840,561 [FREQUENT]** **Is = 59,982,848 [HIGH]** **Forbidden = 28,100** **MONOLINGUALS** **70.6%** The use of synonyms were used (prohibited). Some sentences were completely invented, where origins of words are unknown (e.g. smoking hears everything, cure, they). New words were made from its phonetics (develop, bedantin, bedeni, kievs, others, deliver, featherman). There is evidence of themed words with smoking (cure, develop)	smoking here is prohibited smoking hears what they demand smoking hears the smoking hears everything (2) smoking hears shshs smoking hears deliver smoking hears with others smoky hears th featherman smoking kievs bedantin smoking he s bedeni spoking cure develop	smoking here is prohibited smoking here is not allowed smoking here is permitted smoking here is forbidened smoking kills hetics smoking here is what they do smoking heating is prevented smoking heals with the dead smoking hears what smoking hears is better smoking geirs they do smiking here is smoking hears the bagen smoking hears wihtih yu smoking cures the wedding smoking hears the wedding smoking hears the begger

MULTILINGUALS **70.8%** There is evidence of guesswork from contradictory statements (smoking here is not allowed, smoking here is permitted) from the context. New words created from phonetics (prohibited, hectics, wedding, better). There are more themed words associated with smoking in direct and opposite nature (heals, dead, kills,cures, begger, heating, prevented). There is a lot of evidence of guesswork and the use of context to construct new sentences (smoking cures the wedding, smoking heals with the dead, smoking hears is better, smoking kills hectics), with some benefiting from phonetics.		
WiM_H_H_4 **Fodder was molded and folded inside sand folders** **Number of occurrences in 5.9 billion Corpus** **Fodder = 19,675** **Was = 38,346,000 [HIGH]** **Molded = 3583** **And = 153,492,230 [HIGH]** **Folded = 19,702** **Inside = 789,379 [FREQUENT]** **Sand = 109,730** **Folders = 6338**	fodder was moulded and folded inside sand boulders (6) fodder was modded and bodded inside sand folders fodder was moulded and folded in sight sound folders fodder was molded and boulded like sand folders folder was moulded and folding inside sound holders fooder was moulded and folded inside sand boxes	fodder was moulded inside sand boulders (2) fodders were mouldered inside sand folds fodder was moulded and folded in tight samples poder was modered in sant boders fodder was folded and molded inside sand molders folder was folded and molded inside sand molders fodder was molded and by sand

MONOLINGUALS **82.4%** Phonetic errors were made (boulders,bodded,sound, Folding, holders, soldiers,), New words made from phonetics (like, in, in sight), there were also words created contextually from the phonetics (sand boxes, sam's mind) **MULTILINGUALS** **91.6%** More Phonetic errors were made (boulders, mouldered, folds, modered, boders, sant, molders,.. etc.), More new words were made from phonetics (sample, tight, father, cider, sam's, found, flooding, santsludder, thodor, holes, stormers). Repetition of words was found (sand). There is also evidence of new semantic meaning being created (thought there was a molded folded in the sand folds, father was found the folders, father was molded and holded inside folders, was moulded and folded into sand, father was molded and folded in cider)	thodder was modded and folded in sand solders fodder was modded and moulded inside sand bites fonder was molds inside sam's mind	thdder was molded and fodded inside sand folders thought there was a molded folded in the sand folds father was molded and folded in cider fodder was molded and folded inside sam's folders fodler was moulded and folded inside fodder was moded and in fother was found the folders father is molded in foth scot flooding in wet sket fodder was moulded and boulded in santsulddder folders thodor was folded and molded in red sands and sand holes fodder was modled inside sand stormers father was molded and holded inside folders was moulded and folded into sand

Discussion

This paper aims to systematically combine different levels of cognitive load, linguistic load and contextual meaning in sentences as a framework to predict how well perceived sentences will be in monolinguals and multilinguals and what the perceptual differences are between them in noise.

From the sentence clusters formed in monolinguals and multilinguals, the sentence categories in the theoretical framework played almost no role in which cluster sentences were placed in. As a consequence, one cannot predict how well perceived a sentence will be, or how well perceived one sentence will be from another, by classifying them into different level combinations of linguistic load, cognitive load and contextual meaning and coming to a conclusion, by theory or other means, that all combinations are hierarchically ranked from best perceived to worst. The framework is missing more important factors in order to make an accurate prediction.

What the framework has shown is general trends relating to linguistic load and contextual meaning. Multilinguals performed worse than monolinguals in both speech comprehension and quality of written communication in high linguistic load and no meaning conditions. This is a consequence of cross-language interactions on the phono-lexical and ortho-lexical level as described in the BLINCS model (Shook,2013). This framework has particularly shown how important contextual meaning, and therefore semantic processing, is to speech comprehension in multilinguals. This is because when there is no semantic processing, multilinguals are left with only phono-lexical and ortho-lexical processing that can be easily confused. This study found that there were even instances where entire sentences were remodelled by multilinguals to create new semantic content in order to compensate for the lack of semantic processing. The creation of new words and semantic content from phonetic elements within the original sentence was the biggest factor that caused the difference in the total number of mistakes between monolinguals and multilinguals, which were not considered phonetic errors in the design.

This study found the effects of cognitive load to be miniscule for both monolinguals and multilinguals. This is indicative to the strategies implemented by both groups when writing their answers. It must have been the case that in situations where the sentences made no sense, monolinguals and multilinguals focused on phonetics to try and create semantic meaning that is feasible, rather than result to rote-memorisation, otherwise there would be a more considerable difference.

However, semantic meaning would have less of an impact in monolinguals; monolinguals would have had enough linguistic experience and exposure to recognise a word completely from its phonetics and separate it from other words of similar phonetic construction, with semantic meaning taking a secondary role if necessary. This is why bottom-up models such as Shortlist B (Norris &McQueen, 2008) are more suitable for monolinguals and interactive activation models such as BLINCS are more suitable for multilinguals. For theories like Shortlist B, that proposes word perception to be a probabilistically determined selection of likely candidates influenced by previous words and confirmed through hearing the first phonetic syllables of the word, only speakers with enough linguistic experience could be able to create an accurate list of candidates from previous words for this theory to function well; and the majority of those speakers would be monolinguals. Multilinguals would need to work harder to perceive the sentences well, and an interactive model that allows processing on many levels, not just phonetic, is more appropriate.

The framework has confirmed the appropriateness of certain models over others for monolinguals and multilinguals, with general trends in perception from linguistic load and contextual meaning. However, what are the important factors that are missing from the framework to accurately predict the perception of sentences in noise?

From looking at the sentence clusters and descriptive analyses of sentence outliers, Word Frequency and Sentence Predictability were consistently low throughout the hardest of sentences in both monolinguals and multilinguals, with the ability of a word to be morphed to other words being an extra factor in multilinguals that made them perform much worse than monolinguals in some sentences. When words could be shortened or morphed to form other words (e.g. openness to open or opens) multilinguals performed much worse than monolinguals. The BLINCS model acknowledges lexical frequency to play a role in semantic networks, and easily morphable words have strong semantic and phonetic connections to other similar words that can be easily activated. Monolinguals, on the other hand, were not as sensitive to these words, and it can be explained as having enough linguistic experience to classify similar words differently.

Future research that involves predicting how well perceived a sentence will be in noise for monolinguals and multilinguals should systematically measure word frequency, sentence predictability and morphable words in their framework, as this study shows they take prominence, alongside the effects of contextual meaning and linguistic load. Research should also continue to investigate if Native English Speakers proficient in other languages perform as well as a monolingual in noise and if performance in multilinguals increases the more languages they are proficient in. Factors such as language learning and listening strategies could be an underlying cause for increases in performance in both cases, especially in polyglots (Cohen et al., 2007). Main limitations of this paper include low power in the statistical tests due to the separation of 8 different conditions for a small sample size; and typing speed possibly playing a role in how well written answers were. Despite this however, this study has concluded word frequency, sentence predictability and morphable words to be important measures to predict the perception of sentences in noise for both monolinguals and multilinguals, with linguistic load and context playing important secondary roles, especially in multilinguals, and the length of sentences having no clear effect on speech perception in noise.

References

Bahrick, H. P., Hall, L. K., Goggin, J. P., Bahrick, L. E., & Berger, S. A. (1994). Fifty years of language maintenance and language dominance in bilingual Hispanic immigrants. *Journal of Experimental Psychology: General*, *123*(3), 264.

Bates, E., Devescovi, A., & Wulfeck, B. (2001). Psycholinguistics: A cross-language perspective. *Annual review of psychology*, *52*(1), 369-396.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and cognition*, *12*(1), 3-11.

Bialystok, E. (2011). Reshaping the mind: the benefits of bilingualism. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *65*(4), 229.

Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., & Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. *Journal of Speech, Language, and Hearing Research*, *27*(1), 32-48.

Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *The Journal of the Acoustical Society of America*, *112*(1), 272-284.

Cohen, A. D., & Macaro, E. (2007). *Language learner strategies: Thirty years of research and practice*.

Corpus.byu.edu. (2018, April 1). *NOW Corpus (News on the Web)*. Retrieved from corpus.byu.edu: https://corpus.byu.edu/now/

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1987). Phoneme identification and the lexicon. *Cognitive Psychology*, *19*(2), 141-177.

Cutler, A., Weber, A., Smits, R., & Cooper, N. (2004). Patterns of English phoneme confusions by native and non-native listeners. *The Journal of the Acoustical Society of America*, *116*(6), 3668-3678.

Dijkstra, T., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and cognition*, *5*(3), 175-197.

Flege, J. E., Bohn, O. S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of phonetics*, *25*(4), 437-470.

Flege, J. E., MacKay, I. R., & Piske, T. (2002). Assessing bilingual dominance. *Applied Psycholinguistics*, *23*(4), 567-598.

Florentine, M. (1985). Speech perception in noise by fluent, non-native listeners. *The Journal of the Acoustical Society of America*, *77*(S1), S106-S106.

Florentine, M., Buus, S., Scharf, B., & Canevet, G. (1984). Speech reception thresholds in noise for native and non-native listeners. *The Journal of the Acoustical Society of America*, *75*(S1), S84-S84.

Foster, K. I. (1976). Accessing the mental lexicon. *New approaches to language mechanisms*, 257-287.

Golestani, N., Rosen, S., & Scott, S. K. (2009). Native-language benefit for understanding speech-in-noise: The contribution of semantics. *Bilingualism: Language and Cognition*, *12*(3), 385-392.

Grosjean, F (1988). "Exploring the recognition of guest words in bilingual speech." *Language and cognitive processes* 3.3, 233-274.

Grosjean, F. (1997). Processing mixed language: Issues, findings, and models. *Tutorials in bilingualism: Psycholinguistic perspectives*, 225-254.

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *The Journal of the Acoustical Society of America*, *107*(5), 2711-2724.

Hazan, V., & Simpson, A. (2000). The effect of cue-enhancement on consonant intelligibility in noise: speaker and listener effects. *Language and Speech*, *43*(3), 273-294.

Hedrick, M. S., & Younger, M. S. (2007). Perceptual weighting of stop consonant cues by normal and impaired listeners in reverberation versus noise. *Journal of Speech, Language, and Hearing Research*, *50*(2), 254-269.

Helfer, K. S. (1994). Binaural cues and consonant perception in reverberation and noise. *Journal of Speech, Language, and Hearing Research*, *37*(2), 429-438.

Hockett, C. F. (1955). *A manual of phonology* (No. 11). Waverly Press.

Jia, G., Aaronson, D., & Wu, Y. (2002). Long-term language attainment of bilingual immigrants: Predictive variables and language group differences. *Applied Psycholinguistics*, *23*(4), 599-621.

Jia, G., Strange, W., Wu, Y., Collado, J., & Guan, Q. (2006). Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure. *The Journal of the Acoustical Society of America*, *119*(2), 1118-1130.

Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *The Journal of the Acoustical Society of America*, *61*(5), 1337-1351.

Kaushanskaya, M., Yoo, J., & Marian, V. (2011). The effect of second-language experience on native-language processing. *Vigo international journal of applied linguistics*, *8*, 54.

Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research*, *25*(1), 161-168.

Krizman, J., Marian, V., Shook, A., Skoe, E., & Kraus, N. (2012). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proceedings of the National Academy of Sciences*, *109*(20), 7877-7881.

Krizman, J., Slater, J., Skoe, E., Marian, V., & Kraus, N. (2015). Neural processing of speech in children is influenced by extent of bilingual experience. *Neuroscience letters*, *585*, 48-53.

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, *25*(5), 497-514.

Leather, J., & James, A. (1991). The acquisition of second language speech. *Studies in second language acquisition*, *13*(3), 305-341.

Li, P., & Farkas, I. (2002). 3 A self-organizing connectionist model of bilingual processing. In *Advances in psychology* (Vol. 134, pp. 59-85). North-Holland.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive psychology*, *10*(1), 29-63.

Massaro, D. W. (1989). Testing between the TRACE model and the fuzzy logical model of speech perception. *Cognitive Psychology*, *21*(3), 398-421.

Mayo, L. H., Florentine, M., & Buus, S. (1997). Age of second-language acquisition and perception of speech in noise. *Journal of speech, language, and hearing research*, *40*(3), 686-693.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive psychology*, *18*(1), 1-86.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological review*, *88*(5), 375.

McClelland, J. L., & Rumelhart, D. E. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological review*, *89*(1), 60.

Meador, D., Flege, J. E., & MacKay, I. R. (2000). Factors affecting the recognition of words in a second language. *Bilingualism: Language and Cognition*, *3*(1), 55-67.

Miikkulainen, R. (1993). *Subsymbolic natural language processing: An integrated model of scripts, lexicon, and memory*. MIT press.

Morton, J. (1969). Interaction of information in word recognition. *Psychological review*, *76*(2), 165.

Murdock Jr, B. B. (1962). The serial position effect of free recall. *Journal of experimental psychology*, *64*(5), 482.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*(3), 189-234.

Norris, D., & McQueen, J. M. (2008). Shortlist B: a Bayesian model of continuous speech recognition. *Psychological review*, *115*(2), 357

Norris, D., McQueen, J. M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology*, *34*(3), 191-243.

Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., & Abrams, H. B. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics*, *27*(3), 465-485.

Scharenborg, O., Norris, D., Bosch, L., & McQueen, J. M. (2005). How should a speech recognizer work?. *Cognitive Science*, *29*(6), 867-918.

Shi, L. F. (2009). Normal-hearing English-as-a-second-language listeners' recognition of English words in competing signals. *International Journal of Audiology*, *48*(5), 260-270.

Shi, L. F. (2010). Perception of acoustically degraded sentences in bilingual listeners who differ in age of English acquisition. *Journal of Speech, Language, and Hearing Research*, *53*(4), 821-835.

Shook, A., & Marian, V. (2013). The bilingual language interaction network for comprehension of speech. *Bilingualism: Language and Cognition*, *16*(2), 304-324.

Steeneken, H. J. (2018). *7 Noise Data-base*. Retrieved 3 25, 2018, from ACOUSTEEN: http://www.steeneken.nl/7-noise-data-base/

Takata, Y., & Nábělek, A. K. (1990). English consonant recognition in noise and in reverberation by Japanese and American listeners. *The Journal of the Acoustical Society of America*, *88*(2), 663-666.

The Complete University Guide. (2018). *Durham University*. Retrieved 3 25, 2018, from The Complete University Guide: https://www.thecompleteuniversityguide.co.uk/durham/international

Van Wijngaarden, S. J., Bronkhorst, A. W., Houtgast, T., & Steeneken, H. J. (2004). Using the Speech Transmission Index for predicting non-native speech intelligibility. *The Journal of the Acoustical Society of America*, *115*(3), 1281-1291.

Van Wijngaarden, S. J., Steeneken, H. J., & Houtgast, T. (2002). Quantifying the intelligibility of speech in noise for non-native listeners. *The Journal of the Acoustical Society of America*, *111*(4), 1906-1916.

Weiss, D., & Dempsey, J. J. (2008). Performance of bilingual speakers on the English and Spanish versions of the Hearing in Noise Test (HINT). *Journal of the American Academy of Audiology*, *19*(1), 5-17.

Zhao, X., & Li, P. (2007). Bilingual lexical representation in a self-organizing neural network model. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 29, No. 29).

Zhao, X., & Li, P. (2010). Bilingual lexical interactions in an unsupervised neural network model. *International Journal of Bilingual Education and Bilingualism*, *13*(5), 505-524.

Appendices

APPENDIX A

Marking Rubric for SPERI and SPIn Measures

Mistakes in SPERI were assessed in the fashion of a marking rubric where points were accumulated and 2 final scores presented to reflect the degree of mistakes made. The 2 final scores represented the Total Number of Mistakes Made (including phonetic errors) and the Number of Phonetic Errors Made (separate measure). SPIn measures were dealt individually for each sentence; the binary measure only permits a 0 or 1 to be measured per sentence.

Table A1: *The marking system for SPERI and SPIn.*

SPERI	SPIn
Assign Points to Total Number of Mistakes if:	**Assign a 1 if the 2 points below satisfy:**
• Extra words were added over the total word count of the original sentence (+1 per word)	• If the sentence had the correct subject, verb and object
• Words correctly identified were in the wrong order (+1 per order error)	• If the sentence was capable to be conceptualised with the basic image matching that of the original sentence
• Blanks were left (+1 per word blank)	**Still assign a 1 if:**
• Word Repetition (+1 per word)	• A synonym was used
• Word Mesh or Filler words (e.g. qewhj +1 per filler)	• Words were homophones
• The word had no phonetic resemblance to the original (+2 per word, 1 for getting the word wrong and 1 for not recognising the word in the original sentence in a largely phonetic way)	• The semantic meaning of the sentence was the same
Assign Points to Total Number of Mistakes AND Phonetic Errors if:	• Adjectives were omitted or changed
• The word was phonetically very similar to the original (+1 per word)	**Otherwise issue a 0. A 0 would be given even if:**
• Word was the wrong tense or was plural instead of singular or vice versa (+1 per word)	• Almost all the words were correctly identified but it was missing an essential word for correct semantic meaning (e.g. Bananas created gathered circles of wicked soldiers [Original: Bananas created FROM gathered circles of wicked soldiers], Finder of the fine lines [Original: FIND the fine lines])
Not considered mistakes:	
• If the word was incorrectly spelt (e.g. farmacy for pharmacy)	
• If the words used were homophones to the original words (e.g. which and witch)	
• If the words was spelt phonetically (e.g. SUMBODEE for somebody)	

APPENDIX B

Extract of Questionnaire with definitions of Native and Proficient Language

Questionnaire

For the researcher to complete

Participant ID ________

For the participant to complete

Please fill in or circle the following details:

Sex M / F

Age ______

The native language(s) you speak in this study is considered to be the language(s) you were brought up with (either in school or at home) and practiced consistently from infancy (roughly from before the age of 5) to late teens (to roughly the age of 16 or above). It would be comparable to an educated native speaker of that language.

Should this be difficult to determine, here are some guidelines, otherwise please ignore the bullet points. If it is still unclear, please ask the researcher:

- The term 'native' in this study does not have any connotation to the country you lived in/the ethnicity you are (e.g. If you are Thai, lived your whole life in Thailand but you cannot speak Thai because you went to an international school that had an English curriculum, then English is your native language)
- Your native language(s) should be the languages you are, academically speaking, most confident in.
- If you speak a dialect of an official language and do not speak the official language please write 'Dialect of [Insert official language]' (e.g. Dialect of Mandarin Chinese, Dialect of Hindi), or '[Official Language] including dialects' if you do. If the official language is uncertain, please write down the dialect using the English Alphabet and in brackets the region/country it originates from.

Proficient language(s) are languages that are not your native languages, but you speak them just as well as a native (C1-C2 according to the CEFR levelling system)

If you use your native/proficient language(s) daily, once every other day or at least once a week, please put a ✳ next to it (e.g. English✳)

Native Language(s): ____________________ ____________________

____________________ ____________________

Proficient Language(s):

____________________ ____________________

____________________ ____________________

____________________ ____________________

☐ Pick tick to confirm that you have no known hearing impairments or deafness or difficulty in hearing that you are aware of.

APPENDIX C

Table C1: *The list of sentences used in the experiments, the sentences are in accordance with the conditions imposed upon them.*

Test Trail Sentences	
Test Trail Sentence 1	My friends and I went swimming
Test Trail Sentence 2	That apple is red
Test Trail Sentence 3	Mechanics Fantastic Jolting Fire
Test Trail Sentence 4	Fencing jumpers utility house
Test Trail Sentence 5	Kotolov Yanit Epol
Test Trail Sentence 6	had been to a yuneram before, but I didn´t enjoy the lopticals
Group Name: NoM_L_L	**No Meaning_ Low Linguistic Load _Low Cognitive Load (NoM_L_L)**
NoM_L_L_1	Winter surrounds false width
NoM_L_L_2	Orange batteries promote emptiness
NoM_L_L_3	Wolves distribute excessive listings
NoM_L_L_4	Openness rewards quiet marathons
NoM_L_L_5	Silver chaos enchants poems
NoM_L_L_6	Liquid engines roar anxiously
Group Name: NoM_L_H	**No Meaning_ Low Linguistic Load _High Cognitive Load (NoM_L_H)**
NoM_L_H_1	Extinction of purple corpses occurs from exquisite breath.
NoM_L_H_2	Knowledge amongst pierced rhythms dreamt only of warmth.
NoM_L_H_3	Swollen films of wounded bulbs cleansed exotic lightning
NoM_L_H_4	Agitated persons with reflected, spiky and musical surfaces
NoM_L_H_5	A delivery of underwater tigers enraged spiritual incense
NoM_L_H_6	Bananas created from gathered circles of wicked soldiers
Group Name: NoM_H_L	**No Meaning_ High Linguistic Load _Low Cognitive Load (NoM_H_L)**
NoM_H_L_1	Hail halls healing hell
NoM_H_L_2	Tall coal bowls poll
NoM_H_L_3	Caught cops fought caps.
NoM_H_L_4	Cats pat fat pets
NoM_H_L_5	Bells called billed balls
NoM_H_L_6	Paws clawed thawed straw

Group Name: NoM_H_H	**No Meaning_ High Linguistic Load _High Cognitive Load (NoM_H_H)**
NoM_H_H_1	My rude mood threw blue glue sky high
NoM_H_H_2	My crew blew too few dry white wheats.
NoM_H_H_3	A crowd of clouds bowed their bared hairs
NoM_H_H_4	Bees sue to be by the sea bay
NoM_H_H_5	Ted said red lead rods read seed beads.
NoM_H_H_6	Tanned fans tinned fins then ten spring strings
Group Name: WiM_L_L	**With Meaning_ Low Linguistic Load _Low Cognitive Load (WiM_L_L)**
WiM_L_L_1	Walking to the supermarket
WiM_L_L_2	He ate scrambled eggs
WiM_L_L_3	Smoking here is forbidden
WiM_L_L_4	Your email was received
WiM_L_L_5	We watered the plants
WiM_L_L_6	The pharmacy was closed
Group Name: WiM_L_H	**With Meaning_ Low Linguistic Load _High Cognitive Load (WiM_L_H)**
WiM_L_H_1	He checked his watch to see the time
WiM_L_H_2	We asked for their signatures and shaked hands
WiM_L_H_3	She then decided to put her gloves on.
WiM_L_H_4	They were completely lost, they needed a compass.
WiM_L_H_5	I ordered a delivery, but it never came.
WiM_L_H_6	She went to the store to buy magazines.
Group Name: WiM_H_L	**With Meaning_ High Linguistic Load _Low Cognitive Load (WiM_H_L)**
WiM_H_L_1	He sees pea trees
WiM_H_L_2	Hands in sound sands
WiM_H_L_3	The wared bear stared
WiM_H_L_4	Find the fine lines
WiM_H_L_5	Black rocks blocked locks
WiM_H_L_6	Ducks by thy docks
Group Name: WiM_H_H	**With Meaning_ High Linguistic Load _High Cognitive Load (WiM_H_H)**
WiM_H_H_1	Lice ridden mice hidden in brown round rice
WiM_H_H_2	We write white lies though true truth dies.
WiM_H_H_3	Warm slow storms blow over seesaws on seashores
WiM_H_H_4	Fodder was molded and folded inside sand folders
WiM_H_H_5	Our guests dressed their best wearing western vests
WiM_H_H_6	The night might be bringing stinging frost bites

APPENDIX D

Table D1: *This table shows which clusters each sentence belonged to for monolinguals and multilinguals. Cluster colour states how well perceived the sentence was both in terms of speech comprehension and quality of written responses, sorted hierarchically from worst perceived to best: green>blue>red>yellow>grey>purple. The arrows show if the sentence has moved up [green arrow] or down the hierarchy [red arrow]. The category number corresponds to the definitions given in Table 3 and shows how far the sentence has moved from monolinguals to multilinguals. This table should be read from left to right (e.g. Sentence NoM_L_L_3 was placed in the red cluster for monolinguals but was placed in the blue cluster for multilinguals. The sentence moved down the hierarchy from red to blue [red arrow pointing down]. It moved by a large amount [Category 3])*

	Monolinguals (Fig.7)	Multilinguals (Fig.8)	Monolingual-Multilingual (Fig.9). Numbers show the Category the sentence was placed in
NoM_L_L_1	green	green	1
NoM_L_L_2	red	blue▼	2
NoM_L_L_3	red	blue▼	3
NoM_L_L_4	purple	red▼	3
NoM_L_L_5	grey	red▼	3
NoM_L_L_6	purple	grey▼	2
NoM_L_H_1	red	blue▼	3
NoM_L_H_2	red	blue▼	3
NoM_L_H_3	blue	blue	1
NoM_L_H_4	yellow	red▼	3
NoM_L_H_5	grey	blue▼	3
NoM_L_H_6	yellow	red▼	2
NoM_H_L_1	red	blue▼	2
NoM_H_L_2	red	blue▼	2
NoM_H_L_3	red	green ▼	3

NoM_H_L_4	grey	red ▼	2
NoM_H_L_5	yellow	red ▼	2
NoM_H_L_6	blue	green ▼	1
NoM_H_H_1	red	blue ▼	3
NoM_H_H_2	red	blue ▼	2
NoM_H_H_3	green	blue ▲	1
NoM_H_H_4	green	blue ▲	1
NoM_H_H_5	blue	blue	1
NoM_H_H_6	blue	green ▼	2
WiM_L_L_1	purple	purple	1
WiM_L_L_2	purple	purple	1
WiM_L_L_3	blue	red ▲	1
WiM_L_L_4	purple	purple	1
WiM_L_L_5	purple	grey ▼	2
WiM_L_L_6	purple	purple	1
WiM_L_H_1	purple	grey ▼	2
WiM_L_H_2	purple	grey ▼	2
WiM_L_H_3	purple	grey ▼	2
WiM_L_H_4	purple	purple	1
WiM_L_H_5	purple	purple	1
WiM_L_H_6	purple	purple	1
WiM_H_L_1	purple	purple	1
WiM_H_L_2	green	blue ▲	1
WiM_H_L_3	yellow	red ▼	2
WiM_H_L_4	yellow	red ▼	2
WiM_H_L_5	red	blue ▼	2
WiM_H_L_6	red	blue ▼	3
WiM_H_H_1	red	blue ▼	2

WiM_H_H_2	yellow	red ▼	2
WiM_H_H_3	blue	blue	2
WiM_H_H_4	green	blue ▲	1
WiM_H_H_5	yellow	red ▼	2
WiM_H_H_6	grey	grey	1