Georgian Scientists
ქართველი მეცნიერები
Vol. 7 Issue 2, 2025
https://doi.org/10.52340/gs.2025.07.02.66

Eye-Tracking Emotional and Spiritual Responses in Bilinguals: A Comparative Study of Georgian and Arabic Speakers

Zurab Kherodinashvili; Kholoud Al-Baker

Clinical and Counseling Psychology, Teachers College, Columbia University

Abstract

This study examined how bilingual speakers of Georgian-English and Arabic-English process emotional and spiritual language across their first (L1) and second (L2) languages using eye-tracking methodology. Participants viewed a series of emotionally neutral, negative, and spiritually themed words and sentences presented in both L1 and L2. Eye-tracking measures, including fixation duration, time to first fixation (TTFF), revisit counts, and total gaze time, were analyzed to assess cognitive and emotional engagement with the stimuli. Data from 35 participants (17 Georgian-English and 18 Arabic-English bilinguals) revealed that participants generally exhibited longer fixation durations and greater total gaze time when reading in English (L2), indicating increased cognitive load. However, spiritually themed content elicited longer fixations and higher gaze time in L1, suggesting deeper emotional and cultural resonance. TTFF was faster in L2 for both groups, though Arabic participants showed slower TTFF for spiritual content in Arabic, implying reflective processing. These findings support theories of emotional grounding in L1 and offer implications for multilingual education, therapy, and cross-cultural communication. Cultural and design limitations are acknowledged, with recommendations for incorporating religiosity measures and standardized stimuli in future research.

Key Words: Eye-tracking - Language - Spiritual - Georgian Population - Arab Population - Gaze - Fixation

Recent linguistic research reveals the complexity of language as deeply entangled with its speakers' emotional, psychological, and cultural fabric. This shift in contemporary literature disrupts earlier models of language conceptualization that viewed language through the limited prism of *structural* or *functional* transmission. Structural theories were mainly based on internal linguistic relationships, while functionalist models focused on "speaker-hearer" dynamics and communal needs. However, increasing evidence suggests that both approaches overlook the full depth of how language operates within lived human contexts. Instead, language now appears less as a fixed system and more

as a fluid, dynamic process shaped by emotion. However, such a broad definition of the term makes it almost impossible to create a cohesive scientific narrative.

One of the main challenges that face linguistic research is the multiplicity of the term 'language.' For instance, a wide range of interrelated elements influence the success of language learning and acquisition. These variables include, but are not limited to, age of acquisition, adult-child connections, and overall cultural dynamics (e.g., individualistic vs. collectivist). Therefore, it is unsurprising that language is one of the most extensively researched communication systems, possibly due to the interconnection of many elements (Smith & Kirby, 2008). Due to its conceptual complexity, several scholars from multiple disciplines have tried to broaden the theory of language through their chosen disciplines, including psychology (Miller, 1990) and philosophy (Ekpenyong & Ikegbu, 2018). This being said, a significant chunk of the current understanding of language has been broadened by clinical psychology, offering a new way of understanding how language and language structures may affect human perception and mental health (Lindquist et al., 2015).

A growing body of research within clinical psychology and neuroscience supports the idea that language is not just a tool for expressing emotion but is also integral to how emotions are constructed and understood. The *affective turn* in linguistics is marked by the increased emphasis placed on the role of language in emotion categorization. For instance, Lindquist et al. (2015) suggest that individuals with limited emotional vocabularies may struggle with emotional regulation, identity formation, and interpersonal connection.

Perhaps the most striking example of the limited emotional vocabulary and its role in the emotion regulation hypothesis has been repeatedly demonstrated in psycho-therapeutic settings. Research in clinical psychology has increasingly focused on the unique challenges bilingual individuals face in psychotherapy, especially regarding "language switching"—the transition between a first language (L1) and a second language (L2). One immediate concern involves the emotional engagement/accessibility of early life experiences, with the primary premise being that L1 offers richer emotional engagement/accessibility in therapy than L2. Scholars such as Harris et al., 2006; and Lindquist et al., 2015 have demonstrated that developmental memories encoded in the mother tongue (L1) are more challenging to access and process when therapy is conducted in L2. This "here and now" divide can disrupt emotional regulation and hinder therapeutic progress- particularly in building strong therapeutic alliances.

Further research has indicated a deeper and more intimate link between language and emotion regulation. In particular, researchers such as Nook et al. (2024) and Tang and Ding (2023) have demonstrated how different types of words (concrete and abstract) are processed. In literature, these type words have quite common and heterogeneous definitions, defining concrete type words as words with an apparent physical reference (e.g., an apple) and abstract type words as words without apparent physical reference (e.g., Justice). It should be noted that Abstract words, due to their nature, have become central to studies on emotional regulation. The reasoning behind this interest is inherently linked to similarities between abstract words and emotions, which are anatomically similar. For

example, abstract words and emotions lack clear physical referents and are closely tied to context and subjectivity. Recent work by Nook et al. (2024) found that people who use more abstract words tend to engage in more effective emotion regulation. This finding aligns with the Construal Level Theory (Trope & Liberman, 2010), which is based on the notion that representations of things differ in their degree of abstraction and level of *construal* (subjective interpretation of the stimulus). An example of this notion would be planning to run a marathon in a year (distant abstraction), which leads to a high level of construal (fun, challenging, exciting = more positive regulation) while thinking about running a marathon next week (near=lower abstraction) would lead to low construal (getting the right equipment, correct sleep schedule =more concrete regulation). In other words, the shift to abstract language/increased psychological distance promotes a more flexible range of emotion regulation. However, the exact benefits of abstraction and concreteness, including their neural mechanisms, remain controversial (Nook et al., 2024).

What we do know about the neural mechanisms of abstract and concrete words is that concrete word categories activate sensory and motor brain regions due to their physical characteristics, like shape. For example, as demonstrated by Pulvermüller, 2013 and Cousins et al., 2018 thinking about an "apple" may activate the visual cortex (for its shape), or the gustatory cortex (for its taste).

Abstract words, however, do not activate the brain in the same clear-cut way. Without physical referents, they are believed to rely more on contextual and emotional networks (Moseley et al., 2012). Still, new research challenges the strict division between abstract and concrete word processing. According to Moseley et al. (2012) and Pulvermüller (2013), abstract words can become "concretized" when presented in contexts that align with their meaning, particularly emotional or action-related settings. This suggests abstract words may still activate brain regions linked to emotion and movement. According to the two above-mentioned researches, abstract words carry emotional weight, even without physical form. For example, words like "grief" may activate brain areas involved in facial expressions or bodily sensations associated with sadness (Mortillaro & Dukes, 2018). In this sense, the emotional referent of an abstract word can function similarly to the physical referent of a concrete word in terms of neural activation. In other words, processing abstract emotional words may be closely linked to activation in emotion-based brain regions, reinforcing that language, emotion, and facial expression are deeply interconnected.

Furthermore, the difference between concrete and abstract language, especially when describing emotional events, has been observed in various clinical psychology literature. Nook et al., 2024, indicated that people frequently turn to more abstract language when describing negative emotions, suggesting the role of language in emotion regulation. This leads to the possibility that using abstraction when describing negative (such as traumatic) experiences within a clinical setting may lead to better emotional processing (White & Wild, 2016). While this is a relatively unexplored part of the psycho-linguistic literature, such ideas of linguistic distancing have been observed quite heavily in spiritual publications.

For instance, in meditation literature, the method of linguistic distancing, termed "affect labelling," has been used laboriously (Creswell et al., 2007). Affect labelling is an element of *Mindfulness Meditation* and involves the cognitive effort of vocalizing or "labelling" specific emotions through the body (Torre et al., 2018). Examples of linguistic distancing are also present in religious terminology in various Christian traditions, as demonstrated by spiritual elements such as sermons and prayers (Halbband, 1987, as cited in True, 2023). Similarly, people are more inclined to approach God/Divine/Sacred/Spiritual through dialogue if they engage in internal monologues, often as a regulatory function to resolve interpersonal and internal issues (True, 2023). Additionally, the same research indicates that prayer, like an internal monologue, enables individuals to control their thoughts and actions. However, it is worth mentioning that while in traditional spiritual literature, people define prayer as a tool for self-reflection, clinical research emphasizes the wider psychological advantages of such spiritual practices.

Much like language, *spirituality* is a scientific or phenomenological concept that is as complex as subjective. It defies simple classification and empirical reduction, but it is a crucial component of human individuality and mental health. Despite these insights, empirical investigations comparing spiritual word processing across languages remain scarce. This study addresses this gap by employing eye-tracking technology to assess how Georgian-English and Arabic-English bilinguals process spiritual, negative, and neutral words in their native and second languages. By measuring fixation duration, time to first fixation (TTFF), and gaze total time, we aim to evaluate the depth of emotional and cognitive engagement associated with spiritual vocabulary, thereby attempting to link bilingualism, emotion regulation, and the psycholinguistics of spirituality.

Method

Participants

The study included 38 bilingual participants, evenly divided into two groups: 19 Georgian-English bilinguals and 19 Arabic-English bilinguals. All participants were recruited through Instagram advertisements targeting relevant linguistic and cultural communities. To ensure sufficient English language proficiency, all participants completed the LexTALE test. Only those who scored above the threshold for intermediate-to-advanced proficiency (B1 level) were included in the final sample.

Participants were also screened for eye-tracking data quality. Only those whose data met RealEye.io quality metrics between grades 4 and 6 were retained. These quality grades reflect the accuracy and reliability of the eye-tracking data, with 1 being the lowest and 6 the highest. Two participants from the Georgian-English group were excluded due to missing data in the spiritual word category and low data quality (grade 2). One participant from the Arabic-English group was excluded for similar reasons (grade 3). The final sample, therefore, consisted of 35 participants.

The inclusion criteria for the Georgian group were: aged between 18 and 65, native Georgian speakers with English as a second language, and residing in Georgia. For the Arabic group, the inclusion

criteria were: aged between 18 and 65, native Arabic speakers with English as a second language, and residing in Qatar.

Design

The study employed a within-subjects design with three independent variables: word category (spiritual, negative, neutral), language (L1 vs. L2), and population (Arabic-English bilinguals vs. Georgian-English bilinguals). For each participant, the first language (L1) was either Arabic or Georgian, depending on their population group, and the second language (L2) was English for all participants. All participants were exposed to all levels of the word category and language variables, while population served as a quasi-independent grouping variable. The primary dependent variables were fixation duration (the average length of individual fixations in milliseconds), time to first fixation (TTFF; the time taken to make the first fixation on a target area in milliseconds), and gaze total time (the total gaze duration on a target area in milliseconds). Stimuli were grouped by category on the RealEye.io platform and randomized within each condition to minimize anticipatory effects and prevent implicit learning. Additionally, L1 and L2 stimuli were alternated between the left and right sides of the screen to control for spatial bias. All participants completed the study in one sitting and followed the same sequence of tasks: providing informed consent, completing the LexTALE English proficiency test, participating in the RealEye experimental task, and finally responding to the Adult Attachment Scale and Daily Spiritual Experience Scale.

Materials

RealEye.io for Eye-Tracking Data

Eye-tracking data were collected using RealEye.io (https://www.realeye.io/), a webcam-based platform that employs artificial intelligence, specifically deep neural networks, to estimate eye gaze points. This platform is particularly well-suited for remote data collection, as it does not require specialized hardware such as Gazepoint or Tobii systems. Given the study's focus on basic eye-tracking metrics, including fixation and gaze duration, RealEye.io provided a cost-effective and accessible solution. While future studies may benefit from more advanced tools, this method was appropriate for the current research.

Stimuli consisted of 120 grayscale images, evenly divided into two types: 60 individual word images and 60 sentence images. Word stimuli included 20 spiritual, 20 negative, and 20 neutral words. Sentence stimuli were matched for linguistic complexity and divided equally between L1 and L2 sentences (30 each). All stimuli were randomized in terms of screen position; L1 and L2 items were alternated between the left and right sides of the screen to prevent learning effects and reduce anticipatory bias.

Adult Attachment Scale (AAS)

The Adult Attachment Scale (AAS), developed by Hazan and Shaver (1987), was used to assess attachment styles. It consists of 18 items rated on a 5-point Likert scale and measures three dimensions of attachment: Depend (α = .75), Anxiety (α = .72), and Close (α = .69). Based on participants' scores across these subscales, attachment styles were categorized as Secure, Avoidant, or Anxious.

Daily Spiritual Experience Scale (DSES)

The Daily Spiritual Experience Scale (DSES), developed by Underwood (2011), was used to assess the frequency of daily spiritual experiences, including both theistic and nontheistic items. It has demonstrated high internal consistency (α = .90). Participants responded to 16 items using binary choices (0 = No, 1 = Yes), and total scores were calculated by averaging their responses.

Procedure

This study was approved by the Institutional Review Board at Teachers College, Columbia University. Participants began by reading and agreeing to an informed consent form. They then completed the LexTALE English proficiency test. Those who met the inclusion criteria proceeded to the experimental task, which was conducted using the RealEye.io platform. This task involved viewing pairs of words categorized as spiritual, negative, or neutral, presented in both the participant's native language and in English. Words and sentences were matched for frequency, length, and emotional valence.

Stimuli were randomized within each category. Individual words were presented for approximately 1 to 2 seconds, followed by an interstimulus interval of 1.5 seconds. Sentences were displayed for 3 to 5 seconds. A subset of stimuli included spiritually themed sentences, customized by group, drawn from the Orthodox Bible for Georgian participants and the Islamic Religiosity Measure and the Psychological Measure of Islamic Religiousness (PMIR) for Arabic participants, with English translations. Eye movements were recorded using RealEye.io, with fixation duration, time to first fixation (TTFF), and gaze total time as the primary dependent variables.

Following the experimental task, participants completed two surveys: the Daily Spiritual Experience Scale (DSES) and the Adult Attachment Scale (AAS). The DSES was customized using Orthodox terminology for Georgian participants and Islamic terminology for Arabic participants. The entire procedure was completed in one sitting and took approximately 15 minutes.

Data Analysis

Data were analyzed using Jamovi to assess the effects of word category (spiritual, negative, neutral) and language (L1 vs. L2) on the dependent variables: fixation duration, time to first fixation (TTFF), and gaze total time. A mixed-design ANOVA was conducted, with word category and language as within-subjects factors, and population (Arabic-English vs. Georgian-English bilinguals) as a between-subjects factor.

Prior to analysis, the data were screened for outliers, and participants with low-quality eye-tracking data (below RealEye.io grade 4) were excluded from the analysis. The assumptions of normality and homogeneity of variance were tested using the Shapiro-Wilk test and Levene's test, respectively. No significant violations of these assumptions were found.

Effect sizes were calculated, and the significance level was set at $\alpha = 0.05$. Post-hoc pairwise comparisons were performed using Tukey's HSD test to explore significant interactions or main effects. All statistical analyses were conducted using Jamovi Desktop 2.6.44 (The Jamovi project 2025, *jamovi* Version 2.6).

Results

For the Arab participant group, descriptive statistics for fixation-related measures across language and word categories are presented in Appendix A. For individual words, participants generally exhibited shorter fixation durations, fewer revisits, and lower total gaze time when viewing Arabic (L1) words compared to English (L2) words. Specifically, for spiritual words, the average fixation duration was 291 ms in Arabic and 298 ms in English, while total gaze time was substantially lower in Arabic (2921 ms) than in English (6177 ms) (see Appendix A). The data for spiritual sentences revealed a similar pattern, with nearly identical fixation durations between languages (301 ms for Arabic vs. 302 ms for English), but faster time to first fixation (1136 ms vs. 1346 ms) and reduced total gaze time (1336 ms vs. 1684 ms) in Arabic (see Appendix A). These trends suggest more efficient cognitive processing in participants' first language, particularly for spiritually themed content.

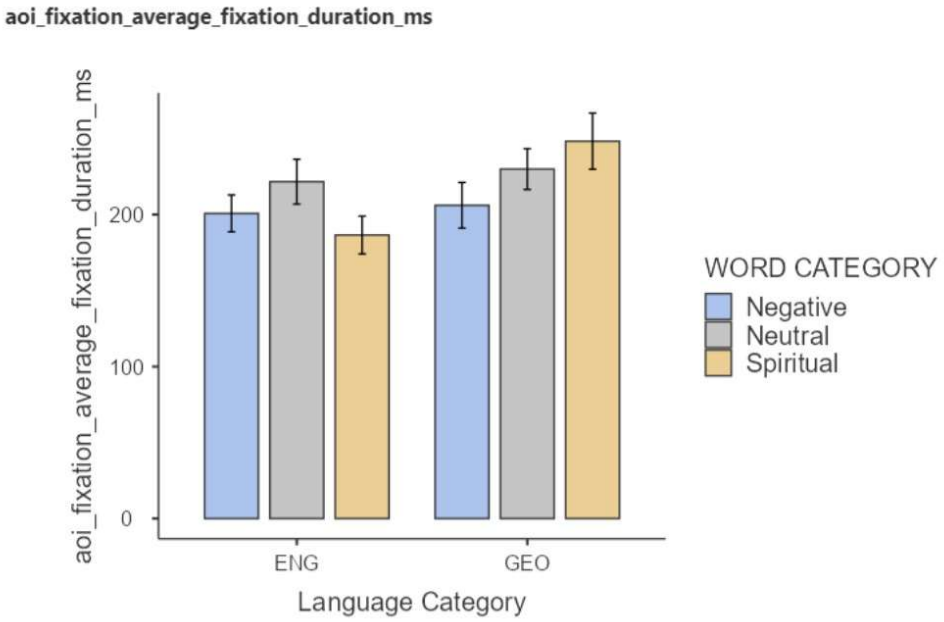
For the Georgian participant group, descriptive statistics for fixation-related measures across language and word categories are presented in Appendix A. For individual words, participants generally exhibited shorter fixation durations, fewer revisits, and lower total gaze time when viewing Georgian (L1) words compared to English (L2) words. For spiritual words, average fixation duration was (243 ms) in GEO and (187 ms) in ENG. The total Gaze time on Georgian Spiritual words were (1271 ms) and (1444 ms) in English translations respectively (See Appendix A). The data for Spiritual Sentences revealed that Georgian participants spent (235 ms) on Georgian Spiritual Sentences and (247 ms) on English Spiritual Sentences. Additionally, Georgian participants had faster fixation times (1233 ms vs 1318 ms) and reduced gaze (1233 vs 1318) in their first language as opposed to their second language (see Appendix A). These trends suggest more efficient cognitive processing in participants' first language, particularly for spiritually themed content.

Fixation Duration

Georgian Population

A two-way ANOVA was conducted to examine the effects of language category (English vs. Georgian) and word category (Negative, Neutral, Spiritual) on average fixation durations in Georgian-English bilinguals. As shown in Bar Chart 1a, when reading in English, participants exhibited the longest average fixation durations for neutral words, followed by negative and then spiritual words. In contrast,

during Georgian reading, spiritual words elicited the longest fixations, followed by neutral and negative words.



Bar Chart 1a. AOI Fixation Duration (ms) for the Georgian Population

ANOVA

ANOVA - aoi_fixation_average_fixation_duration_ms

	Sum of Squares	df	Mean Square	F	p
Word Category	8397	2	4198	3.374	0.038
Language Category	10231	1	10231	8.221	0.005
Word Category ✻ Language Category	2042	2	1021	0.821	0.443
Residuals	141864	114	1244		

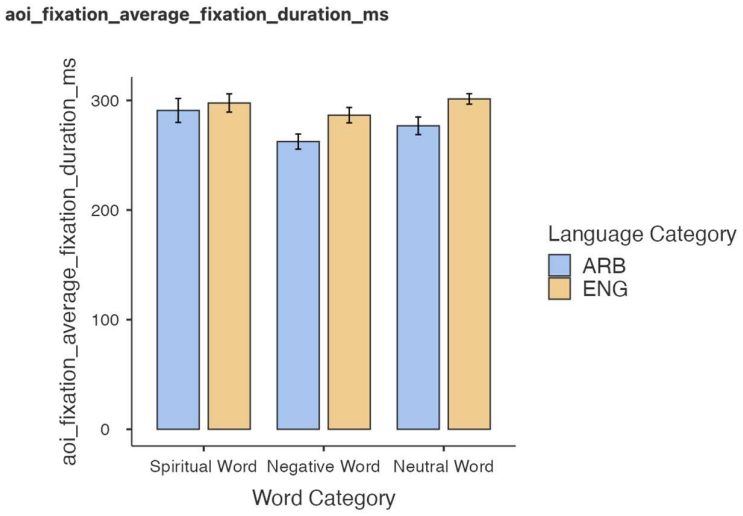
[3]

Table 1a. ANOVA of AOI Fixation Duration (ms) for the Georgian Population

The analysis revealed a significant main effect of Language Category, $F(1, 114) = 4.35$, $p = .039$, indicating that average fixation durations differed significantly between language groups. The main effect of Word Category was not significant, $F(2, 114) = 1.14$, $p = .322$, suggesting no meaningful differences in fixation duration across word types. Additionally, the interaction between Language Category and Word Category did not reach statistical significance, $F(2, 114) = 2.33$, $p = .101$, although the result may suggest a trend worth exploring in future studies. Overall, these findings suggest that language context plays a significant role in modulating visual attention during reading, whereas word category and its interaction with language do not (view Table 1a).

Arab Population

Analysis of fixation duration revealed significant main effects of both Word Category, F(2, 114) = 3.374, p = .038, and Language, F(1, 114) = 8.221, p = .005 (see Table 1b,). These results indicate that participants' gaze behavior varied across spiritual, emotional, and neutral word types, and differed significantly between Arabic (L1) and English (L2). As shown in Bar Chart 1b, participants exhibited longer fixation durations on English words across all word categories compared to Arabic words. Interestingly, greater variability in fixation time by word category was observed in Arabic, with the longest fixations occurring for spiritual words. In contrast, fixation durations in English were relatively stable across categories. This pattern suggests heightened sensitivity to word meaning in the native language, particularly for spiritually salient content.



Bar Chart 1b. AOI Fixation Duration (ms) for the Arab Population

ANOVA

ANOVA - aoi_fixation_average_fixation_duration_ms

	Sum of Squares	df	Mean Square	F	p
Word Category	8397	2	4198	3.374	0.038
Language Category	10231	1	10231	8.221	0.005
Word Category ✻ Language Category	2042	2	1021	0.821	0.443
Residuals	141864	114	1244		

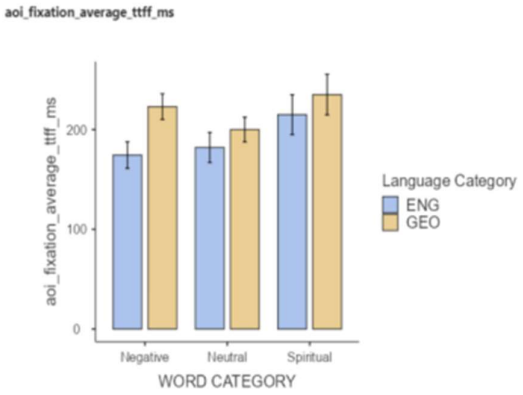
[3]

Table 1b. ANOVA of AOI Fixation Duration (ms) for the Arab Population

Time To First Fixation (TTFF)

Georgian Population

Results for Time to First Fixation (TTFF) indicated a significant main effect of Language, with F(1, 114) = 4.742, p = .031, indicating that language significantly influenced fixation latency. In contrast, WORD CATEGORY showed no significant effect, F(2, 114) = 2.531, p = .084, nor did the interaction between the two factors, F(2, 114) = 0.515, p = .599. These findings suggest that language plays a more prominent role than word type in shaping early visual attention (View Bar Chart 2a; Table 2a).



Bar Chart 2a. Time to First Fixation (TTFF) in Milliseconds (ms) for the Georgian Population

ANOVA

ANOVA - aoi_fixation_average_ttff_ms

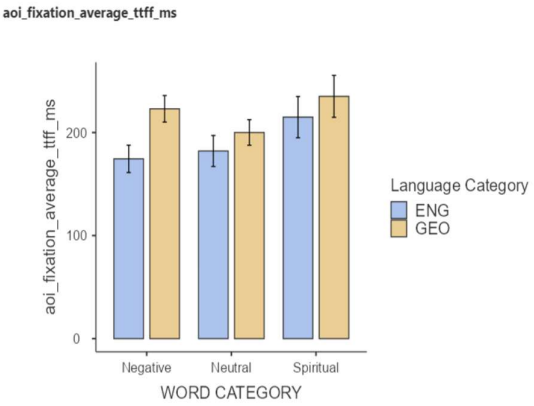
	Sum of Squares	df	Mean Square	F	p
WORD CATEGORY	26267	2	13134	2.531	0.084
Language Category	24606	1	24606	4.742	0.031
WORD CATEGORY ✻ Language Category	5341	2	2671	0.515	0.599
Residuals	591530	114	5189		

[3]

Table 2a. ANOVA of Time To First Fixation (TTFF) in Milliseconds (ms) for the Georgian Population

Arab Population

Results for Time to First Fixation (TTFF) indicated a significant main effect of Language, F(1, 114) = 37.24, p < .001, with participants orienting their gaze faster to words in English (L2) than in Arabic (L1) (see Table 2b). There were no significant effects of Word Category, p = .063, nor a significant interaction between Word Category and Language, p = .984, suggesting that attentional



ANOVA

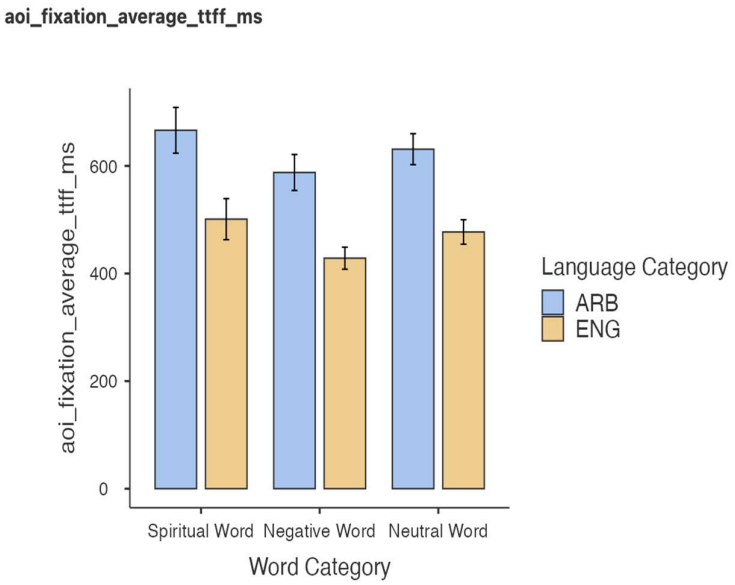
ANOVA - aoi_fixation_average_ttff_ms

	Sum of Squares	df	Mean Square	F	p
WORD CATEGORY	26267	2	13134	2.531	0.084
Language Category	24606	1	24606	4.742	0.031
WORD CATEGORY ✻ Language Category	5341	2	2671	0.515	0.599
Residuals	591530	114	5189		

[3]

orientation occurred at a similar rate across word types in both languages. However, Bar Chart 2b shows that TTFF was consistently faster in English, with minimal variation across word categories. In contrast, Arabic words, especially spiritual ones, elicited slightly longer delays, potentially indicating deeper cognitive or emotional engagement with content presented in the native language. These findings

suggest that while English (L2) may prompt quicker, more automatic gaze orientation, Arabic (L1) may trigger more reflective processing, particularly for spiritually salient content.



Bar Chart 2b. Time to First Fixation (TTFF) (ms) for the Arab Population

ANOVA

ANOVA - aoi_fixation_average_ttff_ms

	Sum of Squares	df	Mean Square	F	p
Language Category	763207	1	763207	37.2399	<.001
Word Category	115809	2	57905	2.8254	0.063
Language Category ✻ Word Category	644	2	322	0.0157	0.984
Residuals	2.34e+6	114	20494		

[3]

Table 2a. ANOVA of Time To First Fixation (TTFF) in Milliseconds (ms) for the Georgian Population

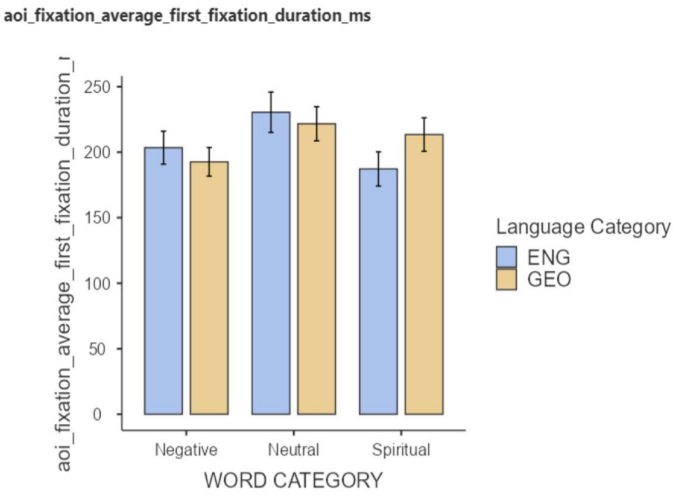
Gaze Total Time

Georgian Population

A two-way ANOVA examining gaze duration (aoi_gaze_average_total_time_spent_ms) revealed a significant main effect of WORD CATEGORY, $F(2, 114) = 3.977$, $p = .021$, indicating that word type influenced sustained visual attention. Language Category showed a marginal effect, $F(1, 114) = 3.734$, $p = .056$, suggesting a possible trend. No significant interaction was found, $F(2, 114) = 0.124$, $p = .884$,

indicating that word effects were consistent across languages. These results suggest differential attentional engagement by word type, relevant for understanding cognitive processing in clinical populations (Table 3a).

First fixation durations were longest for neutral words, shortest for negative words, and intermediate for spiritual words. Fixations were slightly longer in English than Georgian, especially for negative words, indicating faster emotional processing in the native language (Bar Chart 3a).



Bar Chart 3a. AOI Gaze Total Time Spent (ms) for the Georgian Population

ANOVA

ANOVA - aoi_gaze_total_time_spent_ms

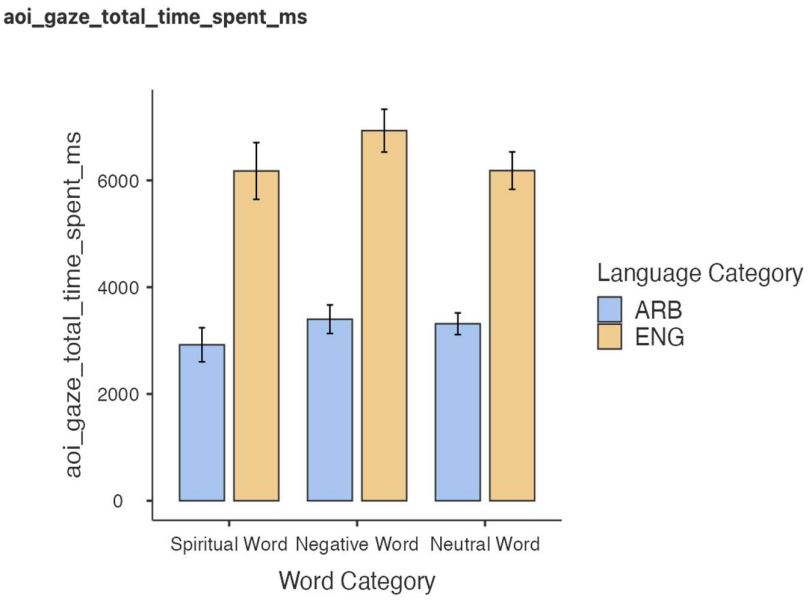
	Sum of Squares	df	Mean Square	F	p
WORD CATEGORY	3.37e+7	2	1.69e+7	33.01	< .001
Language Category	1.63e+6	1	1.63e+6	3.19	0.077
WORD CATEGORY ✻ Language Category	2.72e+6	2	1.36e+6	2.67	0.074
Residuals	5.82e+7	114	510766		

[3]

Table 3a. ANOVA of Gaze Total Time in Milliseconds (ms) for the Georgian Population

Arab Population

The analysis of Gaze Total Time showed a significant main effect of Language, $F(1, 114) = 119.460$, $p < .001$, indicating that participants spent more time looking at words in English (L2) than in Arabic (L1) (see Table 3b). No significant effects were found for Word Category or for the interaction between Word Category and Language, suggesting that the overall amount of visual attention allocated to words was not influenced by their emotional or spiritual content, nor did it vary meaningfully across languages. As illustrated in Bar Chart 3b, gaze time was consistently longer for all word types in English, with negative words drawing the most attention in both languages. This may suggest that English requires more cognitive processing or emotional effort, challenging assumptions that the native language always facilitates deeper engagement.



Bar Chart 3b. AOI Gaze Total Time Spent (ms) for the Arab Population

ANOVA

ANOVA - aoi_gaze_total_time_spent_ms

	Sum of Squares	df	Mean Square	F	p
Language Category	3.11e+8	1	3.11e+8	119.460	<.001
Word Category	7.93e+6	2	3.96e+6	1.523	0.222
Language Category ✻ Word Category	2.23e+6	2	1.11e+6	0.428	0.653
Residuals	2.97e+8	114	2.60e+6		

[3]

Table 3b. ANOVA of Gaze Total Time in Milliseconds (ms) for the Arab Population

Language and Spiritual Words

Georgian Population

A paired samples t-test was conducted to compare fixation durations between language categories and a specific condition labeled as Average Duration SW. The results revealed a highly significant difference, t(11) = –20.8, p < .001, with an extremely large effect size (Cohen's d = –5.99), indicating a substantial difference in fixation duration between the two conditions (Table 4a).

Paired Samples T-Test

Paired Samples T-Test

			statistic	df	p		Effect Size
Language Category	AVERAGE DURATION SW	Student's t	−20.8	11.0	< .001	Cohen's d	−5.99

Note. H_a $\mu_{\text{Measure 1 - Measure 2}} \neq 0$

Normality Test (Shapiro-Wilk)

			W	p
Language Category	-	AVERAGE DURATION SW	0.917	0.264

Note. A low p-value suggests a violation of the assumption of normality

Descriptives

	N	Mean	Median	SD	SE
Language Category	12	1.50	1.50	0.522	0.151
AVERAGE DURATION SW	12	188.68	193.58	31.248	9.021

Table 4a. Paired Samples t-Test Comparing Fixation Duration for Spiritual Words Across Languages for the Georgian Population

Arab Population

A focused t-test comparing fixation duration for spiritual words across languages revealed a significant difference, t(39) = 43.1, p < .001. Participants spent significantly more time fixating on spiritual words in Arabic (L1) than in English (L2), highlighting the greater emotional or cultural resonance of spiritual content when presented in the native language. The validity of this finding is supported by the Shapiro-Wilk test of normality, p = .966, confirming that the data met parametric assumptions. This suggests that spiritually meaningful words in L1 engage participants more deeply, both cognitively and emotionally, than when encountered in L2.

Paired Samples T-Test

Paired Samples T-Test

			statistic	df	p
aoi_fixation_average_fixation_duration_ms	Word Category	Student's t	43.1	39.0	<.001

Note. H_a $\mu_{\text{Measure 1 - Measure 2}} \neq 0$

Normality Test (Shapiro-Wilk)

			W	p
aoi_fixation_average_fixation_duration_ms	-	Word Category	0.989	0.966

Note. A low p-value suggests a violation of the assumption of normality

Table 4b. Paired Samples t-Test Comparing Fixation Duration for Spiritual Words Across Languages for the Arab Population

General Discussion

This study set out to explore how bilingual individuals, specifically Georgian-English and Arabic-English speakers, process emotional and spiritual language across their first (L1) and second (L2) languages. By using eye-tracking technology to measure fixation duration, time to first fixation (TTFF), and total gaze time, we aimed to uncover how language context influences emotional and spiritual engagement. Our findings support the hypothesis that bilingual speakers process emotionally salient and spiritually themed words differently depending on language, with notable cross-linguistic and cultural variations.

Across both populations, participants generally exhibited longer fixation durations and greater total gaze time when engaging with L2 (English) stimuli. This pattern is consistent with cognitive load theory, which posits that processing a second language typically requires more mental resources due to reduced automaticity and emotional grounding. However, this effect was modulated by word type. For example, among Georgian-English bilinguals, spiritual words in Georgian elicited the longest fixation durations, suggesting a deeper level of emotional and cognitive engagement in L1 when content carries personal or cultural resonance. Similarly, Arabic speakers showed significantly longer fixation durations and total gaze time for spiritual content in Arabic, aligning with previous findings in clinical and affective neuroscience that highlight the emotional specificity of L1 in memory encoding and affect regulation.

Interestingly, TTFF results revealed a different pattern. Both groups oriented faster to English (L2) words across categories, perhaps due to the experimental design in which English appeared on alternating sides, or due to the visual salience of unfamiliar words prompting quicker initial fixations. However, Arabic speakers took longer to fixate on spiritual words in Arabic compared to English, potentially reflecting deeper cognitive or emotional appraisal rather than processing speed. This slower initial fixation, coupled with longer gaze durations, supports the idea that spiritually resonant content in L1 triggers more reflective processing, as observed in mindfulness and affects labeling literature.

These findings extend previous research on linguistic embodiment and emotional abstraction. The emotional and spiritual salience of abstract words, like those used in this study, appears to depend heavily on cultural and linguistic context. The longer fixation durations and gaze times for spiritual words in L1 support the hypothesis that such content is processed with more personal significance in one's native language. This has important implications for clinical practice, particularly in multilingual therapeutic settings. Therapists working with bilingual clients should be aware that L2 communication may hinder access to emotionally charged or spiritually meaningful content, potentially impacting therapeutic alliance and emotional regulation.

Notably, both populations exhibited increased engagement with spiritual words compared to neutral and negative words, but only in their respective native languages. This aligns with the concept that spirituality is often embedded in culturally specific language and practices, and that abstract spiritual vocabulary can activate emotion-regulating neural circuits, as shown in recent neuroscientific studies. The use of culturally and religiously tailored stimuli (e.g., Orthodox Christian terms for

Georgians and Islamic phrases for Arabs) likely enhanced this effect, underscoring the importance of cultural congruence in future bilingual research.

Limitations and Future Directions

Cultural Interpretation of Spiritual Content

One limitation of the current study concerns the cultural interpretation of spiritual stimuli. The perception and emotional resonance of spiritual texts may vary significantly within the Georgian population. Cultural or generational norms may influence how individuals engage with emotionally or religiously charged content. For instance, some participants may read spiritual texts quickly out of respect, while others may read them more slowly as a form of contemplation. Similarly, in Arab culture, spiritual words, particularly in Arabic, are often deeply tied to religious practice and everyday language. Some individuals may process these words slowly due to reverence, while others may read them quickly due to high familiarity. These culturally shaped reading behaviors could influence fixation duration and gaze patterns during the task. Future studies should account for individual differences in religiosity and familiarity with spiritual content, potentially using pre-screening surveys or including religiosity as a covariate in analyses.

Social Desirability and Cultural Norms

A further limitation is the potential influence of social desirability and cultural norms on participants' behavior. In more conservative or religious Georgian subgroups, participants may alter their reading behavior, such as reading more slowly or carefully, in response to internalized social expectations or perceived evaluation by researchers. These behaviors may confound dependent measures like total fixation time and time to first fixation. Similarly, Arab participants might feel a cultural or religious obligation to demonstrate respect when reading spiritual words, resulting in prolonged fixations regardless of their emotional engagement. This pressure could lead to attention metrics that reflect social norms more than true cognitive or emotional processing. To mitigate social desirability effects, future research should consider incorporating more implicit measures, such as physiological responses (e.g., pupil dilation), to capture less consciously regulated aspects of emotional engagement.

Stimulus Consistency and Design Limitations

The current study also faced limitations related to stimulus design. Potential incongruities between stimuli across populations, such as differences in visual formatting, font type, or image quality, may have introduced unintended variability in participants' attention. Variations between black-word and white-word stimuli could have affected visual salience or readability, thereby influencing eye-tracking measures. Future studies should ensure that all visual and design elements are standardized across populations to minimize design-based confounds and preserve the validity of cross-linguistic comparisons.

References

Adult Attachment Scale (AAS). (n.d.). Statistics Solutions. Retrieved 17 February 2025, from https://www.statisticssolutions.com/free-resources/directory-of-survey-instruments/adult-attachment-scale-aas/

A Psychological Measure of Islamic Religiousness: Development and Evidence for Reliability and Validity. (n.d.). Retrieved 9 May 2025, from https://www.researchgate.net/publication/248941665_A_Psychological_Measure_of_Islamic_Religiousness_Development_and_Evidence_for_Reliability_and_Validity

Construal-level theory of psychological distance. (n.d.). Retrieved 25 March 2025, from https://psycnet.apa.org/record/2010-06891-005

Cousins, K. A., Ash, S., Irwin, D. J., & Grossman, M. (2017). Dissociable substrates underlie the production of abstract and concrete nouns. *Brain and Language*, *165*, 45–54. https://doi.org/10.1016/j.bandl.2016.11.003

Creswell, J. D., Way, B. M., Eisenberger, N. I., & Lieberman, M. D. (2007). Neural Correlates of Dispositional Mindfulness During Affect Labeling. *Psychosomatic Medicine*, *69*(6), 560–565. https://doi.org/10.1097/PSY.0b013e3180f6171f

*Ekpenyong, B., & Ikegbu, E. A. (2018). The Language of Philosophy and the Philosophy of Language: a Symbiotic Paradigm. Budapest International Research and Critics Institute (BIRCI-Journal) : Humanities and Social Sciences, 1(3), 296–305. https://doi.org/10.33258/birci.v1i3.53*Torre, J. B., & Lieberman, M. D. (2018). Putting Feelings Into Words: Affect Labeling as Implicit Emotion Regulation. *Emotion Review*, *10*(2), 116–124. https://doi.org/10.1177/1754073917742706

El-Dakhs, D. A., Sonbul, S., & Altarriba, J. (2023). How do foreign language learners process L2 emotion words in silent reading? an eye-tracking study. Languages, 8(2), 125. https://doi.org/10.3390/languages8020125

Harris, C. L., Gleason, J. B., & Ayçiçeği, A. (2006). 10. When is a First Language More Emotional? Psychophysiological Evidence from Bilingual Speakers. In A. Pavlenko (Ed.), *Bilingual Minds* (pp. 257–283). Multilingual Matters. https://doi.org/10.21832/9781853598746-012

LexTALE. (n.d.). Retrieved 9 May 2025, from https://www.lextale.com/whatislextale.html

Lindquist, K. A., MacCormack, J. K., & Shablack, H. (2015). The role of language in emotion: Predictions from psychological constructionism. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00444

Miller, G. A. (1990). The Place of Language in a Scientific Psychology. *Psychological Science*, *1*(1), 7–14. https://doi.org/10.1111/j.1467-9280.1990.tb00059.x

Mortillaro, M., & Dukes, D. (2018). Jumping for Joy: The Importance of the Body and of Dynamics in the Expression and Recognition of Positive Emotions. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00763

Moseley, R., Carota, F., Hauk, O., Mohr, B., & Pulvermüller, F. (2012). A Role for the Motor System in Binding Abstract Emotional Meaning. *Cerebral Cortex (New York, NY)*, *22*(7), 1634–1647. https://doi.org/10.1093/cercor/bhr238

Nook, E. C., Ahn, H. E., Schleider, J. L., & Somerville, L. H. (2024). Emotion Regulation is Associated with Increases in Linguistic Measures of Both Psychological Distancing and Abstractness. *Affective Science*, *6*(1), 63–76. https://doi.org/10.1007/s42761-024-00269-7

Pulvermüller, F. (2013). How neurons make meaning: Brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences*, *17*(9), 458–470. https://doi.org/10.1016/j.tics.2013.06.004

Smith, K., & Kirby, S. (2008). Cultural evolution: Implications for understanding the human language faculty and its evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1509), 3591–3603. https://doi.org/10.1098/rstb.2008.0145

Tang, E., & Ding, H. (2023). Emotion effects in Second language processing: Evidence from eye movements in natural sentence reading. Bilingualism: Language and Cognition, 27(3), 460–479. https://doi.org/10.1017/s1366728923000718

THE CHRISTIAN ORTHODOXY SCALE. (n.d.). Retrieved 9 May 2025, from https://scales.arabpsychology.com/s/the-christian-orthodoxy-scale/

The jamovi project (2025). *jamovi* (Version 2.6) [Computer Software]. Retrieved from https://www.jamovi.org

True, R. (2023). Pragmatic Features of Prayer. *Channels*, *8*(1), 81–98. https://doi.org/10.15385/jch.2023.8.1.4

Underwood, L. G., & Teresi, J. A. (2002). The daily spiritual experience scale: Development, theoretical description, reliability, exploratory factor analysis, and preliminary construct validity using health-related data. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine*, *24*(1), 22–33. https://doi.org/10.1207/S15324796ABM2401_04

White, C. N., Kapucu, A., Bruno, D., Rotello, C. M., & Ratcliff, R. (2014). Memory bias for negative emotional words in recognition memory is driven by effects of category membership. Cognition and Emotion, 28(5), 867–880. https://doi.org/10.1080/02699931.2013.858028

White, R., & Wild, J. (2016). "Why" or "How": The Effect of Concrete Versus Abstract Processing on Intrusive Memories Following Analogue Trauma. *Behavior Therapy*, *47*(3), 404–415. https://doi.org/10.1016/j.beth.2016.02.004

Zlatev, J., & Blomberg, J. (2015). Language may indeed influence thought. *Frontiers in Psychology*, *6*, 1631. https://doi.org/10.3389/fpsyg.2015.01631

Appendices

Appendix A

Georgian Population

Table 1a

Descriptive Statistics of Word Categories

Descriptives

	WORD CATEGORY	Language Category	aoi_fixation_average_ttff_ms	aoi_fixation_average_fixation_duration_ms	aoi_gaze_total_time_spent_ms
Mean	Negative	ENG	174	201	2528
		GEO	223	206	2885
	Neutral	ENG	182	222	1822
		GEO	200	230	2343
	Spiritual	ENG	215	187	1444
		GEO	235	243	1271
	Spiritual Sentences	ENG	1318	247	12389
		GEO	1233	235	13980
Standard deviation	Negative	ENG	56.3	51.5	810
		GEO	51.5	60.1	947
	Neutral	ENG	70.4	69.2	661
		GEO	60.7	65.8	804
	Spiritual	ENG	89.5	55.6	570
		GEO	90.9	79.7	424
	Spiritual Sentences	ENG	441	28.6	3785
		GEO	371	28.3	5096
Shapiro-Wilk W	Negative	ENG	0.946	0.881	0.943
		GEO	0.986	0.955	0.968
	Neutral	ENG	0.964	0.937	0.907
		GEO	0.936	0.976	0.963
	Spiritual	ENG	0.988	0.917	0.956
		GEO	0.985	0.947	0.936
	Spiritual Sentences	ENG	0.957	0.984	0.973
		GEO	0.964	0.950	0.976
Shapiro-Wilk p	Negative	ENG	0.367	0.027	0.327
		GEO	0.994	0.576	0.807
	Neutral	ENG	0.584	0.170	0.041
		GEO	0.130	0.823	0.502
	Spiritual	ENG	0.995	0.087	0.473
		GEO	0.984	0.321	0.201
	Spiritual Sentences	ENG	0.264	0.918	0.619
		GEO	0.391	0.173	0.708

Table 2a

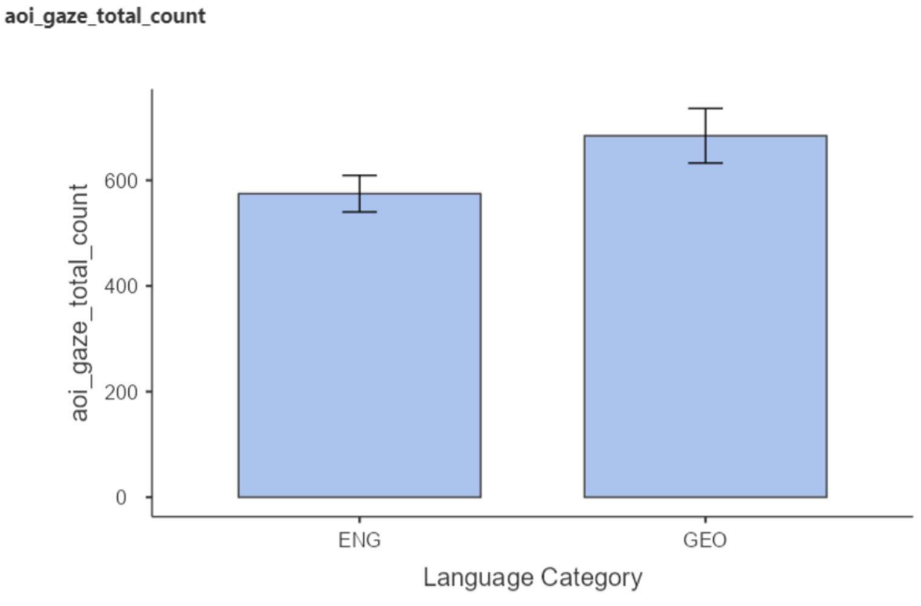
Descriptive Statistics of Spiritual Sentences

Descriptives

	WORD CATEGORY	Language Category	aoi_fixation_average_ttff_ms	aoi_fixation_average_fixation_duration_ms
N	Spiritual Sentences	ENG	30	30
		GEO	30	30
Missing		ENG	0	0
		GEO	0	0
Mean		ENG	1318	247
		GEO	1233	235
Median		ENG	1275	247
		GEO	1191	226
Standard deviation		ENG	441	28.6
		GEO	371	28.3
Minimum		ENG	690	183
		GEO	568	183
Maximum		ENG	2251	321
		GEO	2321	283

Bar Graph 4

Spiritual Sentence Comparison (Total Gaze Count)



Arab Population

Table 1b

Descriptive Statistics of Word Categories

Descriptives

Descriptives

	Word Category	Language Category	aoi_fixation_average_fixation_duration_ms	aoi_fixation_average_first_fixation_duration_ms	aoi_fixation_average_ttff_ms	aoi_revisit_total_count	aoi_gaze_total_time_spent_ms
Mean	Spiritual Word	ARB	291	301	666	0.650	2921
		ENG	298	311	501	0.950	6177
	Negative Word	ARB	262	282	588	0.750	3400
		ENG	287	308	429	2.45	6933
	Neutral Word	ARB	277	300	631	0.750	3315
		ENG	301	326	477	2.15	6184
Standard deviation	Spiritual Word	ARB	49.0	45.0	190	0.745	1428
		ENG	37.2	41.0	171	0.887	2376
	Negative Word	ARB	30.9	33.2	150	0.550	1197
		ENG	31.4	35.6	91.5	1.39	1795
	Neutral Word	ARB	35.9	40.3	129	0.851	909
		ENG	21.4	22.5	101	1.27	1565

Table 2b

Descriptive Statistics for Spiritual Sentences

Descriptives

Descriptives

	Language Category	Word Category	aoi_fixation_average_fixation_duration_ms	aoi_fixation_average_ttff_ms	aoi_gaze_average_total_time_spent_ms
N	ARB	Spiritual Sentence	30	30	30
	ENG		30	30	30
Mean	ARB		301	1136	1336
	ENG		302	1346	1684
Standard deviation	ARB		16.5	271	363
	ENG		16.1	264	325