Revisiting lexical processing: Evidence from Greek-speaking adults

Maria Katsiperi
Aristotle University of Thessaloniki
katsiperim@enl.auth.gr

Eleni Fleva
Aristotle University of Thessaloniki
eleni.fleva@googlemail.com

Abstract
In the present study we examined factors that affect word recognition in Greek while further exploring whether the process is influenced by participants’ educational background. We employed an on-line lexical decision task to a group of 60 young monolingual Greek-speaking adults. Participants had difficulty recognising real words of low frequency while those with lower vs. higher academic attainment were even more vulnerable to this effect. Word length and gender effects were detected but to a smaller extent. The results demonstrate that frequency is the most important factor in word identification although the role of extra-linguistic factors is not to be ignored.

Keywords: word recognition, frequency, length, gender, educational background

1. Introduction
A long-standing debate in the psycholinguistic literature concerns the structure of the mental lexicon and the factors affecting lexical access. Factors argued to be involved in the activation of lexical units are related both to word properties and to the reader’s characteristics. From early on, word frequency has been shown to be one of the most crucial factors in lexical access (Forster & Chambers 1973; Marslen-Wilson 1987; Savin 1963 among others; see Monsell 1991; Murray & Forster 2004 for reviews) with frequently encountered words being more easily recognisable than less commonly used words. Whaley (1978) proposed that frequency is the most significant factor in word recognition while frequency seems to affect not only visual but also spoken word recognition (Dahan, Magnuson & Tanenhaus 2001). Carreiras, Álvarez and de Vega (1993) found that even frequency of the syllables may affect word recognition. Apart from frequency, there are a number of other factors that are relevant to the process. Length of the word, measured either by letters, syllables or phonemes, also plays an important role with shorter words eliciting shorter reaction times (Gough 1972; Whaley 1978; Yap & Balota 2009). However, not all studies have found a length effect (Henderson 1982; Weekes 1997). This may be due to the fact
that shorter words are usually more frequent; thus, length may be confounded with frequency effects. Other factors have been also found to interact in the process. Age-of-acquisition, i.e. at what age a person has learned a given word, and neighborhood size, the number of words that can be created by changing one letter of a target word, have been also reported to crucially interact in the process (Andrews 1997; Barry et al. 2001, Coltheart et al. 1977; Morrison & Ellis 2000). However, they will not be discussed in the present study as they are not part of the experimental design.

Apart from clearly linguistic factors, lexical accessing is also related to individual differences. During the past years, the umbrella term General Language Ability was used to explain differences in language proficiency among L1 speakers (Bowey 1990). Recent research has tried to discriminate the components of this term. The role of orthographic processing skills, defined as the ability to form, store and assess orthographic representations (Bråten et al. 1999; Burt 2006 for a review), of phonological processing skills, the ability to identify and manipulate units of oral language (Unsworth & Pexman 2003) and of print exposure, the degree of investment a person dedicates in reading and literacy activities (Stanovich & West 1989), have been pointed out. These skills, which are inter-related yet distinct, seem to facilitate word recognition. Increased print exposure is associated with faster responses to both words and non-words. Participants with higher level of print exposure are found to be faster and more accurate in naming pseudowords than those with lower level of print exposure (Chateau & Jared 2000). Findings from the English Lexicon Project (Balota et al. 2007; http://elexicon.wustl.edu), an online behavioral database providing results from 816 participants in a lexical decision task, showed that higher vocabulary knowledge was positively correlated with faster and more accurate word recognition (Yap et al. 2012) and faster nonword identification (Yap et al. 2015). In a more general framework, recent studies have detected differences on language tasks as a function of language proficiency either measured by standardised tests (Pakulak & Neville 2010) or by academic achievement (Dabrowska & Street 2006). The findings were attributed to an extended familiarity with written language. These results indicate that language experience is a factor that can differentiate performance on language interpretation in general and on word recognition in particular.
2. The present study

The present study aims to examine to which extent some of the aforementioned factors influence visual word recognition among Greek native speakers/readers. We investigated the frequency and length factor along with grammatical gender. The educational background factor was also investigated. Our purpose is not only to discover whether or not and to which extent the aforementioned factors influence the word identification process in Greek but also to examine how they interact.

2.1 Participants

A group of 60 young monolingual Greek-speaking adults participated in the study ($M = 23.01$, $SD = 3.5$). The group of the participants was divided into two subgroups: (a) thirty (30) university students and (b) thirty (30) age-matched adults, attending post-secondary, college studies. The educational background categories were formulated based on the educational level participants attended: tertiary vs. post-secondary. For the tertiary education level, we included students of Humanities & Science Departments and for the post-secondary level, we recruited participants from Speech Therapy & Physiotherapy programs. We intended to examine participants of both literature (Humanities Department) or language (Speech Therapy) oriented studies as well as Science and Technological (Physiotherapy) majors in order to neutralise a possible advantage of literature or language students due to their further practice with written language. Their number was counterbalanced across categories.

2.2 Experimental Design and Material

The research team created an on-line word recognition experiment in which participants were visually presented with letter strings in lower case. The experiment was developed and displayed using E-prime 2.0 (Schneider, Eschman & Zuccolotto 2002). Participants were instructed to decide as fast as possible whether the given item was an existing word or not, while trying to keep the error rate low. Items were presented on a computer screen and participants answered by pressing the pre-specified buttons for existing/non-existing words. Material was presented in random order. The measures of interest were (a) accuracy and (b) speed of response.

The task includes 140 items in total, 60 existing, 60 pseudo and 20 illegal words in Modern Greek. All the existing words were taken from the GreekLex corpus (Ktori, van Heuven & Pitchford 2008). Pseudowords were created based on existing words.
beginning with a cluster. Subsequently, the first or the second consonant was replaced by another one while respecting the phonotactic rules of the language. For example, the existing word “σκάλα” (skála = ladder) turned into the meaningless but phonotactically legal item “σλάλα” (slála) after the substitution of its second consonant. The same process was followed for the construction of the illegal items.

Existing words beginning with a cluster were modified by changing one of the initial consonants but in this set of words the phonotactic rules were violated resulting into new, illegal words. For instance, the word “ζβέρκος” (zvérkos = nape) was used for the construction of “ζρέκος” (zrékos) which is phonotactically illegal in Greek.

All the words were inanimate nouns. The non-existing items were derived from inanimate nouns as well. No items were used twice, that is to say, the existing words of the task were not used in any case as the base of the non-existing or of the illegal ones.

Frequency, number of syllables and gender suffixes were controlled for. Existing words and pseudowords were divided into three blocks of frequency: high, mid and low. Frequency of the existing words was taken from the GreekLex corpus (Ktori, van Heuven & Pitchford 2008). It was reported as the number of occurrences of the word per million. In the present task, they were lemma frequencies that were taken into account. High-frequency words ranged from 550 to 250 lemma frequency per million words, mid-frequency words varied from 150 to 50 and for low-frequency words the range was from 1 to ≈ 0. A gap between the categories was kept for the blocks to have a quantitative as well as a qualitative difference in frequency between them. Each frequency block contained 20 items.

The frequency of the pseudowords was controlled through their initial cluster. We used the frequencies of the tokens in word-initial position, calculated per million and converted to decimals, as a measure of frequency. All values were taken from the Protopapas corpus for consonants (Protopapas et al. 2010; Institute for Language & Speech Processing (ILSP): PsychoLinguistic Resource). Pseudowords were of high (6.2 to 1), mid (0.8 to 0.1) and low (0.07 to 0.01) frequency with each block containing 20 items. Clusters illegal in initial position do not have frequency, thus illegal words could not be controlled for frequency.

The items were also categorised according to grammatical gender. Only masculine and feminine endings were used. From a total of 140 items, 70 of them carried a masculine suffix and 70 items a feminine one. Masculine and feminine items were
counterbalanced across the frequency blocks. The choice of the suffixes was made based on their prototypicality in terms of predictive value. The feminine endings in use were –η and –α, whose predictive values for feminine gender assignment are 0.93 and 0.82 respectively (Mastropavlou & Tsimpili 2011). The masculine endings were –ας and –ος, which gain a predictive value for masculine gender assignment of 0.83 and 0.77 correspondingly (Mastropavlou & Tsimpili 2011).

For the word-length variable, Two-syllable and three-syllable items were employed, counterbalanced across item word status, frequency and gender categories. The complete list of the items used is displayed in Appendix A.

2.3 Data Analysis

The measures of interest were the accuracy (accurate or non-accurate identification of an item as a word or a non-word) and speed of response (reaction time) in milliseconds. Reaction times (RTs) give us insight into the processing of the items, with longer RTs being associated with greater difficulty in processing. Only the RTs of correctly identified items were included in the statistical analysis. Extreme values (RTs above 5000 msec) were excluded. Outlier RTs, defined as 2 standard deviations above or below each participant’s mean RT per condition, were substituted by the mean in the respective category. This procedure affected less than 1% of the data.

3. Results

Results on the accuracy rates are illustrated below followed by results on the RTs. Regarding accuracy, we report the findings in terms of (a) word status and (b) frequency within each word status category. There was no differentiation as a function of grammatical gender or length and thus we do not comment on these two factors. Turning to the processing data (RTs), differences were detected for all variables and are presented below.

3.1 Accuracy

To explore differences as a function of word status, we compared the accuracy rates of existing, pseudo and illegal words.

---

1 The suffix with the higher predictive value for masculine assignment is –ης (0.92) but it was avoided as it is closely correlated with animacy, a factor that was excluded from the present task.
A one-way ANOVA analysis with word status as the independent variable and accuracy percentages as the dependent one, revealed a significant word status effect ($F_{2,479} = 36.2, p = .000$). Participants demonstrated a ceiling performance in recognising pseudo- and illegal words (95% and 99% respectively) but accuracy rates dropped on real words (86.7%). Bonferroni post-hoc comparisons showed that participants scored significantly lower on the real words compared to both pseudo- ($p = .000$) and illegal ($p = .000$) words. Illegal words significantly outperformed pseudowords as well ($p = .04$).

For the frequency variable, we performed a Univariate analysis with accuracy percentages as the dependent variable and word status along with frequency as the fixed factors. As illegal words do not carry any frequency value, they were excluded from the analysis.

<table>
<thead>
<tr>
<th>Word Status</th>
<th>existing</th>
<th>pseudo</th>
<th>illegal</th>
</tr>
</thead>
<tbody>
<tr>
<td>86.7</td>
<td>95.2</td>
<td>99</td>
<td></td>
</tr>
</tbody>
</table>

*Table 1: Accuracy rates as a function of word status*

<table>
<thead>
<tr>
<th>Word Status</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>existing</td>
<td>high</td>
</tr>
<tr>
<td></td>
<td>mid</td>
</tr>
<tr>
<td></td>
<td>low</td>
</tr>
<tr>
<td>pseudo</td>
<td>high</td>
</tr>
<tr>
<td></td>
<td>mid</td>
</tr>
<tr>
<td></td>
<td>low</td>
</tr>
</tbody>
</table>

*Table 2: Accuracy rates as a function of frequency by word status*

There was a significant word status by frequency interaction ($F_{2,354} = 202.8, p = .000$). To explore it, a one-way ANOVA analysis was conducted for each word status category separately. A frequency effect was manifested for both real ($F_{2,177} = 278; p = .000$) and pseudowords ($F_{2,177} = 3.3, p = .03$). Among real words, high and mid-frequency items outperformed the low-frequency ones (99% vs. 61.8%, $p = .000$ and 99.3 vs. 61.8% respectively). No significant difference was found between high and mid-frequency words. Turning to pseudowords, the reverse pattern was detected.
High-frequency items had significantly lower accuracy scores than mid-frequency ones (94% vs. 96.7, \( p = .03 \)), that is they were misconceived as real ones significantly more. The differences between the high and the low-frequency pseudowords as well as between the mid- and the low-frequency categories were not significant.

3.2 Reaction times

Participant’s RTs on each word status category are presented in Table 3.

<table>
<thead>
<tr>
<th>Word Status</th>
<th>existing</th>
<th>Pseudo</th>
<th>illegal</th>
</tr>
</thead>
<tbody>
<tr>
<td>950</td>
<td>1305</td>
<td>1093</td>
<td></td>
</tr>
</tbody>
</table>

*Table 3: Mean RTs as a function of word status*

A one-way ANOVA analysis with word status as the independent variable and RTs as the dependent one revealed a main effect of word status (\( F_{2, 179} = 20.8, p = .000 \)). Subsequent Bonferroni comparisons demonstrated that participants needed significantly more time to decide on pseudowords compared to both real (\( M = 1305 \) vs. \( M = 950, p = .000 \)) and illegal words (\( M = 1305 \) vs. \( M = 1093, p = .001 \)). Real words were identified significantly faster compared to both pseudowords and illegal items (\( M = 950 \) vs. \( M = 1093, p = .03 \)).

To detect frequency effects, a Univariate analysis was employed, with word status and frequency category as the independent variables and RTs as the dependent one.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>existing</th>
<th>pseudo</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>834</td>
<td>1544</td>
</tr>
<tr>
<td>mid</td>
<td>855</td>
<td>1263</td>
</tr>
<tr>
<td>low</td>
<td>1266</td>
<td>1232</td>
</tr>
</tbody>
</table>

*Table 4: Mean RTs as a function of frequency by word status*

The results revealed a significant word status effect (\( F_{1,354} = 30.2, p = .000 \)), a marginally significant frequency effect (\( F_{2,354} = 2.9, p = .055 \)) and a significant interaction between the two variables (\( F_{2,354} = 10.8, p = .000 \)).
analysis was performed on each word status condition. Among real words, a frequency effect was manifested \((F_{2,179} = 65.7, p = .000)\) with Bonferroni post-hoc tests showing that low-frequency words triggered significantly longer RTs than both high \((p = .000)\) and mid-frequency words \((p = .000)\). The reverse pattern was exhibited among pseudowords. High frequency pseudowords elicited longer RTs than mid \((M = 1544 \text{ vs. } M = 1263)\) and low frequency items \((M = 1263 \text{ vs. } M = 1232)\) with differences however not being significant.

For the word length variable, we performed a Univariate analysis with word status and number of syllables as fixed factors and RTs the dependent one.

<table>
<thead>
<tr>
<th>Length</th>
<th>Word Status</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>existing</td>
<td>pseudo</td>
<td>illegal</td>
<td></td>
</tr>
<tr>
<td>2 syllables</td>
<td>940</td>
<td>1238</td>
<td>1019</td>
<td></td>
</tr>
<tr>
<td>3 syllables</td>
<td>1028</td>
<td>1330</td>
<td>1155</td>
<td></td>
</tr>
</tbody>
</table>

*Table 5: Mean RTs as a function of length by word status*

There were significant word status \((F_{2,354} = 35.1, p = .000)\) and length effects \((F_{1,354} = 12.3, p = .001)\) but not an interaction between them \((F_{2,354} = .255, p = .775)\). However, in order to explore the trend, we conducted independent-samples t-tests in each word status category. A length effect was manifested among real and illegal items \((t(118) = -2.7, p = .006; t(118) = -2.2, p = .02)\) but not among pseudowords.

Finally, to explore possible differences on the RTs as a function of gender, a Univariate analysis with word status and gender as the independent factors and RTs as the dependent ones was conducted.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Word Status</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>existing</td>
<td>pseudo</td>
<td>illegal</td>
<td></td>
</tr>
<tr>
<td>masculine</td>
<td>966</td>
<td>1332</td>
<td>1074</td>
<td></td>
</tr>
<tr>
<td>feminine</td>
<td>1036</td>
<td>1254</td>
<td>1099</td>
<td></td>
</tr>
</tbody>
</table>

*Table 6: Mean RTs as a function of gender by frequency*

The analysis verified a main effect of word status \((F_{2,354} = 32.4, p = .000)\) but no main effect of gender \((F_{1,354} = .03, p = .850)\) or an interaction between the two
variables ($F_{2,354} = 2.06, p = .128$) occurred. Still, in order to detect the differences detected in the means, we performed exploratory independent-samples $t$-tests within word status category. A gender effect was detected for the existing words ($t(118) = -2.6, p = .04$) with masculine words marking significantly shorter RTs than feminine words (t(118) = .354, $p = .04$, $M = 966$ vs. $M = 1036$). The reverse pattern was manifested for the pseudowords with the difference however not being significant ($M = 1332$ vs. $M = 1254$). RTs of the illegal words were almost equal for the masculine and feminine items ($M = 1074$ vs. $M = 1099$).

3.3 Educational background

As mentioned before, lexical properties are not the only source of variation in word recognition. Readers’ characteristics may also influence the process. In the following analysis we have split the sample with respect to their educational background (tertiary vs. post-secondary education) and we explored whether the two groups are differentiated. We report results on both accuracy and RTs for the word status and frequency factor.

3.3.1 Accuracy

To begin with, we performed a Univariate analysis with word status and educational background as fixed factors and accuracy rates as the dependent one.

<table>
<thead>
<tr>
<th>Education</th>
<th>existing</th>
<th>pseudo</th>
<th>illegal</th>
</tr>
</thead>
<tbody>
<tr>
<td>tertiary</td>
<td>90</td>
<td>95.7</td>
<td>99</td>
</tr>
<tr>
<td>post-secondary</td>
<td>83.5</td>
<td>94.7</td>
<td>98.9</td>
</tr>
</tbody>
</table>

Table 7: Accuracy rates for the two educational level groups as a function of word status

There was a main effect of word status ($F_{2,174} = 92.37, p = .000$) and of education type ($F_{1,174} = 11.4, p = .001$) with a significant interaction ($F_{2,174}= 176.7, p = .001$). Independent-samples $t$-tests showed that there was a significant difference as a function of educational background for existing words; students attending tertiary education outperformed those attending post-secondary studies ($p = .04$).
Adding the frequency factor in the analysis and conducting a Univariate ANOVA revealed main effects of word status ($F_{1,348} = 113.3$, $p = .001$), of frequency ($F_{2,348} = 246.4$, $p = .001$) and of educational background ($F_{1,348} = 16.7$, $p = .001$) with a significant interaction between all three variables ($F_{2,348} = 8.6$, $p = .001$).

<table>
<thead>
<tr>
<th>Word Status</th>
<th>existing</th>
<th></th>
<th>pseudo</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Education</strong></td>
<td><strong>high</strong></td>
<td><strong>mid</strong></td>
<td><strong>low</strong></td>
<td><strong>high</strong></td>
</tr>
<tr>
<td>tertiary</td>
<td>99.5</td>
<td>99.8</td>
<td>68.7</td>
<td>94.3</td>
</tr>
<tr>
<td>post-secondary</td>
<td>98.5</td>
<td>98.7</td>
<td>54</td>
<td>93.7</td>
</tr>
</tbody>
</table>

*Table 8: Accuracy rates for the two educational level groups as a function of frequency by word status*

Paired-sample $t$-tests into each frequency group showed that the students of tertiary education answered significantly more accurately than their peers of post-secondary studies, on mid- and low-frequency real words (mid: $M = 99.8$ vs. $M = 98.7$, $t(58) = 2.2$, $p = .04$; low: $M = 68.7$ vs. $M = 54$, $t(58) = 3.6$, $p = .001$).

3.3.2 Reaction Times (RTs)

Turning to the RTs, a Univariate ANOVA with word status and educational background as the fixed factors showed that the time participants needed to answer was differentiated as a function of word status ($F_{2,174} = 22$, $p = .001$) but independent of educational background as no main effect or interaction occurred.

<table>
<thead>
<tr>
<th>Word Status</th>
<th>existing</th>
<th>pseudo</th>
<th>illegal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Education</strong></td>
<td><strong>existing</strong></td>
<td><strong>pseudo</strong></td>
<td><strong>illegal</strong></td>
</tr>
<tr>
<td>tertiary</td>
<td>937</td>
<td>1208</td>
<td>1100</td>
</tr>
<tr>
<td>post-secondary</td>
<td>965</td>
<td>1415</td>
<td>1084</td>
</tr>
</tbody>
</table>

*Table 9: Mean RTs for the two educational level groups as a function of word status*

Adding the frequency factor, apart from the expecting word status and frequency effect ($F_{1,348} = 31.8$, $p = .001$; $F_{2,174} = 22$, $p = .001$), an education type effect was also manifested ($F_{1,348} = 3$, $p = .009$) but there was no interaction between the variables ($F_{2,348}=1.27$, $p = .280$).
Dividing the data into frequency blocks for each word status group revealed that university and college students had significant differences among real words of low frequency ($M = 1175$ vs. $M = 1370$, $t(58) = -2.2$, $p = .03$) and pseudowords of mid frequency ($M = 1178$ vs. $M = 1360$, $t(44) = -1.6$, $p = .02$) with the former answering faster in both cases. The same pattern applied across all categories; the differences however were not significant.

### 4. Discussion

In the present study, 60 monolingual Greek-speaking participants performed a lexical decision task. The lexical variables that we investigated were these of word status (existing vs. pseudowords vs. illegal words), frequency (high vs. mid vs. low), length (2 vs. 3 syllables) and gender (masculine vs. feminine). The role of participants’ educational background was also attested (tertiary vs. post-secondary). Summarising the findings, a word status effect was manifested. Existing words elicited the shortest RTs, that is, participants could quickly decide on their lexicality. Pseudowords triggered the longest RTs while illegal words were in the middle. The difference between pseudo and illegal words is in accordance with previous, well-established findings reporting greater difficulty in rejecting more plausible words than less plausible ones (Coltheart et al. 1977). Regarding accuracy, performance on existing words was poorer than on pseudo and illegal words. Performance dropped even more for the post-secondary students. This finding was explained once frequency was taken into account. Existing words of low frequency were more probable to be identified as non-existing words in comparison to words of high and mid-frequency. They also elicited longer RTs. The mirror image occurred for the pseudowords. High frequency items were more probable to be identified as real words, as previously reported in the
study of Yap et al. (2015). Participants attending post-secondary education were even more vulnerable to this effect. It is thus shown that frequency has a robust effect on the word recognition process while its magnitude also depends on the reader’s characteristics. Coming to the length variable, whose effects were questioned in the literature (Henderson 1982; Weekes 1997), we report significant differences for real and illegal words between bi-syllabic items and tri-syllabic ones. The former were identified significantly faster than the latter. Processing of pseudowords was not affected by length. It could be that the puzzling nature of pseudowords, in that they look like real words but they are not, that masked any effects. Finally, grammatical gender affected RTs when interacting with word status. Existing masculine words were identified faster than feminine words and the reverse pattern applied for pseudowords. Items carrying masculine endings elicited significantly longer RTs than the ones with feminine endings. The fact that the gender effect was stronger among pseudowords can be explained by previous studies reporting that individuals rely on suffix gender cues only in the absence of other lexical information (Hohlfeld 2006 for German, Mastropavlou & Tsimpli 2011 for Greek). The weaker lexical bias of the feminine ending in comparison to the masculine one could stem from the phonologically double gender status of the feminine ending -η. The phoneme i corresponds to both feminine and neuter endings. It seems that, even though only the feminine orthographic representation was employed, there could be a partial activation of the neuter gender that gave rise to the gender effect.

Finally, regarding the educational background factor, the evidence shows that it does differentiate participants’ performance. Participants attending post-secondary education are more affected by frequency than participants in tertiary education and that was manifested both on accuracy rates and on RTs. The reason behind this pattern is not yet clear. The already established fact that skilled readers read faster (Martin 2004) is not enough to explain the finding as it could not justify neither the drop on accuracy rates neither the influence of frequency. Another explanation could be that less experience readers have just not encountered some of the words of low frequency. This may be possible but it cannot justify the effect on processing (RTs) or on pseudowords. We suggest that the educational background effect stems from differences in exposure and familiarity to print language. We argue that extended experience with printed language can boost the orthographic decoding system,
promote higher metalinguistic awareness and give an advantage that is not limited only to the already known forms but expands to novel ones as well.

5. Conclusion
The findings from the on-line word recognition task indicated how some lexical and domain-general factors affect word identification in Greek. In accordance to the literature, frequency was proved to have a great impact on the process (Monsell 1991; Murray & Forster 2004 for reviews). Length and gender were also found to affect it. The role of language experience, measured by educational level, was designated. It remains for future research to examine potential effects of other variables, such as neighborhood size in Greek, as well as to further examine the role of language experience in the word recognition process.

**Acknowledgment:** The study is part of the project *Individual Differences in Anaphora Resolution: Linguistic and Cognitive Effects* (Excellence I, Action 2656, PI: Prof. Ianthi Maria Tsimpli) run at the Aristotle University of Thessaloniki, Greece. It was co-financed by the European Union (European Social Fund - ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework - NSRF).

**References**


Appendix A

Examples

Existing words (60 items)

High frequency
Masculine: ρόλος (role), αγώνας (struggle)
Feminine: πλευρά (side), ἕνωση (connection)

Mid frequency
Masculine: μύθος (myth), αστέρας (star)
Feminine: πίστη (faith), οθόνη (screen)

Low frequency
Masculine: πορθμός (strait), μάλωπας (bruise)
Feminine: κράση (temperament), ανέμη (swift)

Pseudowords (60 items)

High frequency
Masculine: θροχός (τροχός-wheel), σφελετός (σκέλετός-skeleton)
Feminine: γράση (δράση-action), γλίμακα (κλίμακα-scale)

Mid frequency
Masculine: φιλάδος (κλάδος-field), βγαζηήρας (βραστήρας-kettle)
Feminine: σλάλα (σκάλα-ladder), φλεγάτα (φρεγάτα-frigate)

Low frequency
Masculine: κβουνός (κρουνός-faucet), κνίβανος (κλίβανος-sterilizer)
Feminine: χθένα (χτένα-comb), χναμύδα (χλαμύδα-toga)

Illegal (20 items)
Masculine: λκορός (σκόρος-moth) νρατήρας (κρατήρας-crater)
Feminine: λτοντή (βρωμία-dirt), σπατάλη (σπατάλη-waste)