

### **3. Children's Reading of Irish-English Homographs: The Role of Language Background and Task Context<sup>1</sup>**

***Fiona Lyddy<sup>2</sup>, Bryan Roche and Helen Ambrose***

*This study examined how language background and context affect children's reading of Irish-English homographs, words with shared spelling but with different pronunciations and meanings in the two languages (for example, 'fear', 'bean' and 'teach'). Children from Gaeltacht and non-Gaeltacht areas completed a computerised task which involved categorising presented stimuli as 'Irish', 'English', 'Both' or 'Neither'. The language of the task was also manipulated (Irish or English). The results reveal an advantage for Gaeltacht groups on Irish words, particularly within an Irish task context. The non-Gaeltacht group did significantly better in recognising that the homographs could be 'both', although detecting the ambiguity may not be advantageous for biliteracy. Analysis of errors showed effects of task context and the language background of the child, with implications for both reading instruction and assessment.*

#### **Introduction**

For the bilingual learner to read more than one alphabetic language, the acquisition of multiple print-sound (grapheme-phoneme) correspondences is generally required. Such exposure enhances both knowledge of print representations (Bialystok, 1997) and phonological awareness skills. In addition, between language facilitation has been demonstrated in Spanish-

---

1 Acknowledgements: The Gaeltacht section of this research was funded by Cumas Teo. Ré Marnell provided invaluable assistance with translation of materials and data collection. The authors gratefully acknowledge this support.

2 Address for correspondence: Dr. Fiona Lyddy, Dept. of Psychology, National University of Ireland, Maynooth, Co. Kildare. Email: [fiona.lyddy@nuim.ie](mailto:fiona.lyddy@nuim.ie)

#### **42 Children's Reading of Irish-English Homographs**

---

English (Ganschow and Sparks, 1995), French-English (Comeau, Cormier, Grandmaison and Lacrois, 1999), Italian-English (D'Angiulli, Siegal and Serra, 2001) and even Chinese-English (Wang, Perfetti and Liu, 2005). This latter finding is particularly striking given that, in Chinese orthography (which is logographic rather than alphabetic) graphemes map onto syllabic morphemes, not phonemes. Several studies suggest positive effects on English reading of first learning another alphabetic script (e.g. Lesaux and Siegal, 2003; Spencer and Hanley, 2003) and the influence may extend to spoken language as well as literacy. Phoneme awareness emerges only *after* we are taught to read (Goswami, 2002) and becomes a key predictor of reading success as well as language learning. Thus, language affects literacy and literacy affects language, with additional cross-language influences in the case of the bilingual.

The cited positive effects of biliteracy are particularly striking given that most alphabetic language pairs have print-sound mappings that conflict. This is well illustrated by the existence of cross-language homographs. Homographs are polysemous words, words which are lexically identical (i.e. share spelling) but have different meanings; for example, in English, the word 'bank' can be used as a noun meaning 'financial institution' or 'riverside', or as a verb meaning to 'deposit', 'tilt' or to 'rely (on)'. While frequency affects comprehension of the word, disambiguation depends to a large extent on the context in which the word is encountered. In English, homographs can have different pronunciations as well as different meanings; for example, 'row', 'read', 'wind' and 'tear' each have two pronunciations. The process of mapping print to sound, a critical aspect of reading referred to as phonological recoding (e.g., see Ziegler and Goswami, 2005), is complicated given such forms, which reflect the depth of orthography (Frost, Katz and Bentin, 1987), that is, the degree of regularity in the representation of spoken sounds by written letters, in a given language.

Homographs also exist across languages, bringing further implications for the bilingual who must acquire a deep-orthography language such as English. Many such forms will share meaning while pronunciations differ; for example, the French/English 'table' or the Irish/English 'cat' refer to

the same entity. By contrast, non-cognate interlexical homographs are words which are lexically identical in two languages but have different meanings and generally different pronunciations. For example, 'pain' means bread in French, 'Kind' means child in German and 'fear' means man in Irish; in each case the pronunciation differs from the English. These 'false friends' often represent a stumbling block for the language learner, but are seldom problematic for competent bilinguals. Readers seem able to initiate the necessary mappings for the currently-active language without interference from the other language such that ambiguous items are often not noticed. The currently active language mode (see Grosjean, 2001) is key; Gerard and Scarborough (1989), using Spanish-English homographs, found that word frequency in the currently active language, rather than the overall frequency of use in the two languages, predicted recognition of homographs. Thus, tasks that require subjects to identify a homograph are difficult, and children, lacking mature metalinguistic awareness, should have particular difficulty in acknowledging the ambiguity in a word.

All things being equal therefore, if Irish-English bilinguals are presented with a homograph such as TEACH - a written form that occurs in both Irish (as a noun meaning 'house') and in English (as a verb meaning 'to instruct'), with different pronunciations and meanings - it should be read as appropriate to the currently active language mode. Presented in the absence of context, or with both language modes active, it should be read by frequency; here the noun TEACH, meaning 'house' is more frequent in written Irish than is the verb in written English, and therefore the more frequent interpretation would win out. However, for Irish-English processing, the pervasive influence of English with respect to both spoken and written media may place the Irish language mode at an immediate disadvantage, even for those whose everyday language is Irish. This effect might in particular be worth exploring in young readers who are just acquiring competence in both (written) languages and will encounter many such homographs among their early words. Examples include: *air* (on-him), *bean* (woman), *fear* (man), *lean* (follow), *long* (ship), *rang* (class), *sin* (that), *teach* (house).

#### 44 *Children's Reading of Irish-English Homographs*

---

This study therefore aims to examine children's ability to categorise letter sequences as being Irish, English, homograph words or nonsense words, as a function of two factors. Firstly, we examine the effect of the language context in which the task is presented, that is, whether the task itself is encountered in Irish or English. Secondly, the language background of the participants may be influential. The participants were children aged 11-12 years from Gaeltacht and non-Gaeltacht areas, with the Gaeltacht group further broken down by the language that was primarily spoken at home, that is, whether the children reported speaking mainly Irish or English when at home. The effects of these factors, language background and task context, as well as their interactions, are examined with regard to homograph reading in particular.

### **Method**

Stimulus lists were constructed with stimuli of each of four types: Irish words, English words, homographs and nonwords. All consisted of 3-5 letter sequences, of 1-2 syllables. English words (e.g. face, must, out) were selected using the Kucera-Francis (1967) written frequencies. The majority of these words would have been familiar to children, with a few more difficult items included. The Irish word list (e.g. liom, beir, rud) was then matched on an item-by-item basis to the English words, controlling the number of letters and syllables as well as written frequency using *Corpas Náisiúnta na Gaeilge* (The National Corpus of Irish) (ITÉ, 2003). Only the base forms of Irish words (i.e. non-mutated forms) written without a *síneadh fada* (vowel length marker) were selected. The homograph word list (words in both Irish and English e.g. TEACH, LEAN, AIR) was matched to the Irish and English lists for number of letters and syllables. These stimuli were biased towards Irish, that is, written frequencies for the Irish interpretation were higher than for the English interpretation. The final set of stimuli used pronounceable nonwords (e.g. SGEL, ELTE, HOV), which were selected to match the word lists by number of letters and syllables.

Using nonword recognition as a basic index of reading skill, participants who scored above a criterion of 50% were selected for data analysis. (Poorer performance would suggest a difficulty with written word

discrimination or poor adherence to the task instructions.) This produced a sample of 78 children; 41 living and attending school within the Conamara Gaeltacht and 37 living and attending an English-medium school in a Leinster town. The Gaeltacht group had two sub-groups; for one group Irish was the main language outside school and for the other English was the main language outside school. This gave three language background groups.

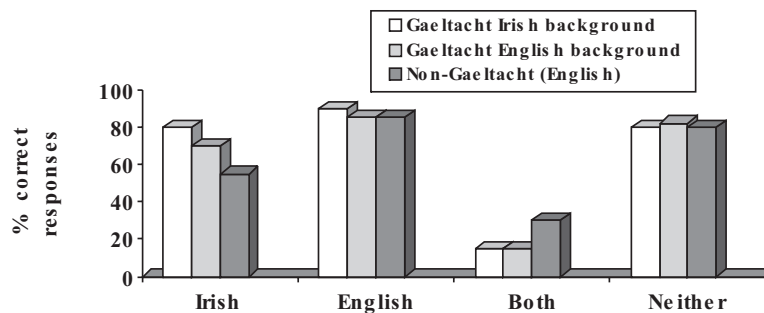
Across these, the language context of the task was manipulated, with random allocation to conditions. That is, children either completed the task through Irish or through English. This involved communication regarding the task (via computer instructions or the researchers' interactions) being conducted through one language, although the Irish context for the non-Gaeltacht group was not complete. That is, the Irish language manipulation differed for the Gaeltacht and non-Gaeltacht groups as the standard of Irish employed could not be the same in both cases. Therefore the non-Gaeltacht group received a simplified version of the context manipulation using materials similar to those used in their Irish class (these children could not have been expected to follow the on-screen instructions for the computer task in Irish whereas the Gaeltacht children had no difficulty with these). Researchers for each group were from, or worked in, the local area, and their language use reflected this.

The stimuli were presented via a computerised lexical decision task. They appeared one at a time, in random order, with each item remaining on the screen until the child pressed the appropriate response key. For each item, the child categorised the word as 'Irish', 'English', 'both' or 'neither' by pressing one of four corresponding response keys on the computer. (The child received pre-training in the use of these before the task proper, which also highlighted the concept of homographs and exposed the child to some examples.) Each response was recorded by the computer program, which computed the number of correct responses for each child for each of the four stimulus types.

## Results

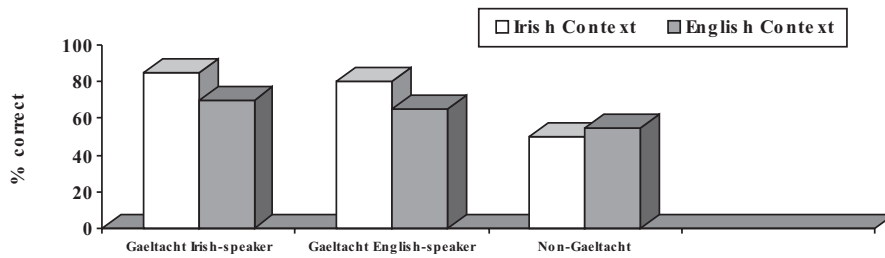
Details of inferential statistics are omitted here but stated effects are statistically significant. Figure 1 shows the mean number of correct responses for the four stimulus types by language background. A straight comparison of performance across the three word types reveals that, for all language groups, performance was better for English words. However the parity of the word lists might well be questioned here: while matching was attempted based on frequency and word/syllable length, the lists may not compare on many key psycholinguistic measures, since data on these are not available for the Irish language. Performances on the English words and the nonsense words are similar across the three language backgrounds and statistical analysis confirmed no differences between the groups on these measures. Differences between the groups are apparent, however, on the Irish word list and the homograph stimuli, and it is here that context effects are also found.

**Figure 1. Correct responses for each stimulus type for the three language background groups**



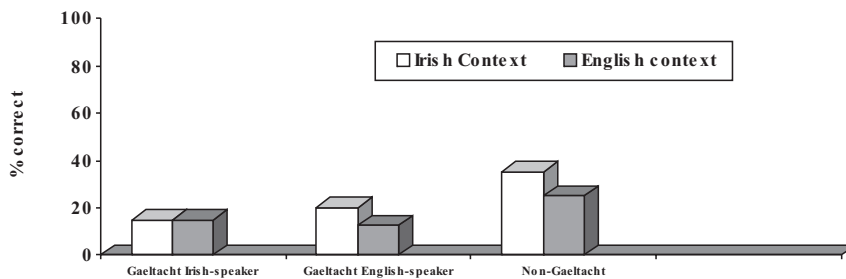
The two Gaeltacht groups show a significant advantage for Irish words over the non-Gaeltacht group; no statistically significant difference was found between scores of the Gaeltacht Irish-speakers and the Gaeltacht English-speakers. Both Gaeltacht groups' advantage on Irish words is augmented in the Irish task context, as illustrated in Figure 2. For the non-Gaeltacht, English-schooled sample, performance on Irish words does not differ significantly as a function of context.

**Figure 2. Correct responses for Irish words for the three language background groups by task context**



By contrast, the non-Gaeltacht group performed significantly better than the two Gaeltacht groups on the homographs. Performance on the ‘both’ stimuli was the poorest of all the stimulus types, for all language groups (see Figure 1 above), as might be expected given the difficulty of such a task, particularly for children. Looking at the number of correct responses here (i.e. recognising that such a stimulus could be ‘both’ Irish and English) a clear advantage is seen for the non-Gaeltacht children, who produce twice as many correct responses. However, recognition that such items are ambiguous is not necessarily advantageous for biliteracy attainment. Looking at Figure 3, while it may appear that the Irish context produces better performance for the English-speaking groups, this effect does not reach statistical significance.

**Figure 3. Correctly detected homographs by language background and context**

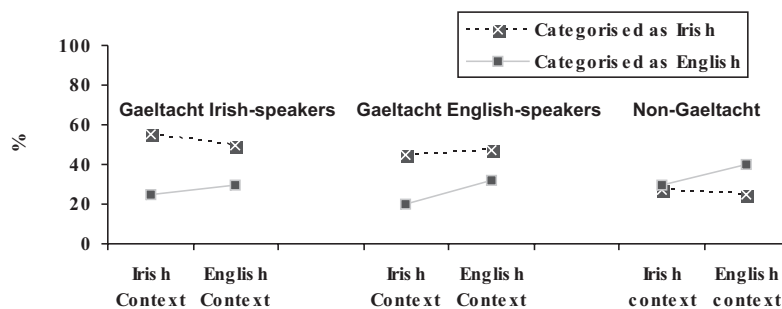


#### 48 Children's Reading of Irish-English Homographs

Of particular interest here is the type of error made on these stimuli, as a function of the language background of the child and the language context in which the task was presented. If a child is more likely to respond to a 'both' stimulus (e.g. TEACH) as being Irish, for example, this would suggest that reading the word automatically activates the Irish representation for this word; in other words this reflects the child's current primary processing language. It would be of interest to know the extent to which this is affected by the language background of the child and the language context in which the task is presented.

Errors made were therefore classified by whether the child responded to the stimulus as Irish or English (i.e. given TEACH for example, does the child respond by pressing the key for 'Irish' or 'English', where the correct response key here is that associated with a 'both' response, to indicate that the written word TEACH is both an Irish word and an English word). It can be seen that overall the children from both Gaeltacht backgrounds were far more likely to err by 'seeing' the words as Irish than as English, while the non-Gaeltacht children (who got more of the homograph responses correct to begin with) produced more English errors. Overall, significantly more errors of the English type occurred within the English task context while context did not affect the Irish type errors. The pattern is illustrated in Figure 4.

**Figure 4. Errors made on the homograph words by language background and context**



## **Conclusions**

Compared to the non-Gaeltacht group, the Gaeltacht groups performed better on the Irish words, and did as well on the English words. It is notable that this advantage was augmented in the Irish task context. This finding has implications for comparable tasks and testing situations, where performance might be maximised by adopting a favourable language mode.

While the pattern of performance suggested superior performance of Irish versus English speakers from the Gaeltacht, any differences were not statistically significant. The children were assigned to these groups on the basis of self-report regarding the main language spoken at home, and it is possible that more stringent criteria in this respect could have identified more differences here. Alternatively, more sensitive materials may have augmented any real differences. The pattern of performance on the homograph stimuli is particularly informative. Compared to the non-Gaeltacht group, the Gaeltacht groups make more errors on the homograph task and are more likely to err by interpreting the words as Irish. This pattern is not necessarily detrimental, as it may reflect a processing mode that avoids capturing the ambiguity in the stimulus, a necessary prerequisite for fluent reading here. Interestingly, the tendency to interpret the words as Irish occurs in both the Irish and the English context, suggesting that these Gaeltacht children continue to operate within an Irish language mode or to interpret the stimuli by dominant frequency, which favours Irish. By contrast, the effect of the switch to the English context can be seen in the non-Gaeltacht group, where the error pattern changes towards English, compared to an undifferentiated pattern in the Irish context. These data suggest that the Gaeltacht-schooled children are not as susceptible to the English context as might be feared given the pervasive influence of English, and the written language in particular.

It might, however, be considered salient that, of the three word types, performance on the English words was best: all three groups performed best on these words, with no differences between language groups. This may reflect a real advantage for English words, for all language groups. However, it could also be an artefact of stimulus selection, rather than a true reflection of the children's word knowledge. The psycholinguistic

resources available for the Irish language are sparse compared to those available in English and utilised here, and, as a result, the extent of matching achieved between the Irish and English stimuli was constrained. Furthermore, contrasts between the languages mean that matching could only be approximated; in particular word type was not the same in items across categories. Ideally the lists might have been based on frequency-matching alone; but here we had to eliminate items with a *sineadh fada* or a mutated form, as these would be identified as Irish words based on those cues alone. Likewise, simply translating items is problematic for word recognition tasks, as the length of the word can affect recognition, or provide cues to one language over another via syllable structure. Further data collection would therefore be required before we can explain why, in all groups, performance was superior for English words.

The health of the language-literacy relationship is important, not just for the individual's literacy attainment but also for the survival and well being of a language. The current research is encouraging with respect to the status of written Irish for Gaeltacht-schooled children, but it also highlights the influence of language mode or context, in interaction with language background, in a word recognition task. The implications for reading and spelling instruction and assessment may well extend beyond simple word recognition.

## **References**

- Bialystok, E. (1997). Effects of bilingualism and biliteracy on children's emerging concepts of print. *Developmental Psychology*, 33, 3, 429-440 .
- Comeau, L., Cormier, P., Grandmaison, E., and Lacrois, D. (1999). A longitudinal study of phonological processing skills in children learning to read a second language. *Journal of Educational Psychology*, 91, 1, 29-43.
- D'Angiulli, A., Siegal, L., and Serra, E. (2001). The development of reading in English and Italian in bilingual children. *Applied Psycholinguistics*, 22, 479-507.
- Frost, R., Katz, L., and Bentin, S. (1987). Strategies for visual word recognition and orthographic depth: A multilingual comparison.

- Journal of Experimental Psychology: Human Perception and Performance*, 13, 104-115.
- Ganschow, L. and Sparks, R. (1995). Effects of direct instruction in Spanish phonology on the native-language skills and foreign-language aptitude of at-risk foreign language learners. *Journal of Learning Disabilities*, 28, 2, 107-120.
- Gerard, L. and Scarborough, D. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 2, 305-315.
- Goswami, U. (2002). Phonology, reading development, and dyslexia: A cross-linguistic perspective. *Annals of Dyslexia*, 52, 141-163.
- Grosjean, F. (2001). The bilingual's language modes. In Janet L. Nicol (ed.) *One Mind, Two Languages: Bilingual Language Processing*. Oxford: Blackwell.
- Institiúid Teangeolaíochta Éireann (2003). *Corpas Náisiúnta na Gaeilge/ National Corpus of Irish*, Volume 1. Dublin: ITÉ.
- Kucera, H. and Francis, W.N. (1967). *Computational Analysis of Present-Day American English*. Providence, Rhode Island: Brown University Press.
- Lesaux, N.K. and Siegel L.S. (2003). The development of reading in children who speak English as a second language. *Developmental Psychology*, 39, 6, 1005-1019 .
- Spencer, L.H. and Hanley, J.R. (2003). Effects of orthographic transparency on reading and phoneme awareness in children learning to read in Wales. *British Journal of Psychology*, 94, 1-28.
- Wang, M., Perfetti, C.A. and Liu, Y. (2005). Chinese-English biliteracy acquisition: cross-language and writing system transfer. *Cognition*, 97, 67-88.
- Ziegler, J.C. and Goswami, U. (2005). Reading acquisition, developmental dyslexia and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, 131, 1, 3-29.