Automagic Pilot

Summary Report: Process evaluation of the Automagic Welsh language intervention pilot conducted September 2023 to July 2024

Author: Peter Russell

Automagic Pilot

Summary Report: Process evaluation of the Automagic Welsh language intervention pilot conducted September 2023 to July 2024

Author: Peter Russell

Research supervisor: Dr Sion Jones (Senior Lecturer, Cardiff University School of Social Sciences)

Executive Summary of Findings

- The Automagic intervention appears to be effective at teaching the intervention content and recombinant 'chunking' skills to a minority of learners (20% scoring 8/10 or higher). The lack of progress in the majority of learners appears to largely be due to high levels of disengagement.

- Knowledge of Welsh Language Continuum (WLC) content was poor across all institutions. It is unclear if this is due to low levels of content retention, or variations in coverage. Some teachers drew attention to the effects of the Covid pandemic as a significant factor in the Welsh language performance of learners. Intervention schools scored slightly higher in WLC assessments, but this accounted for only 3% of the variation in scores.

- Learners reported a relatively high level of enjoyment of WLC lessons, with the majority of learners reporting some or high enjoyment. In interviews, many cited the use of game-based learning as a key reason for this positive disposition.

- No association was found between learner enjoyment or perceived importance of Welsh and WLC assessment scores. These would usually be strong indicators of success. Comparatively, a significant association was found between the intervention scores and reported enjoyment of Welsh.

- WLC content knowledge was often found to be formulaic, limiting learners' ability to utilise language patterns in a flexible manner to facilitate communication.

- Participants who used the intervention were significantly more likely to believe they were making progress in their Welsh language ability (even when high performance was not evident in the assessment data).

- Lower levels of motivation and intervention assessment scores were associated with learners who self-identified as 'non-Welsh'.

- In interviews many learners expressed a perceived lack of utility in learning the Welsh language, often identifying it as a purely academic pursuit confined to the school context.

- Intervention use appears to polarise learner self-efficacy, with intervention participants more likely to believe themselves to be 'very good' or 'very poor' compared to those in the control group.

- In the absence of standardised assessment, both learners and teachers drew on localised comparative measures of Welsh language competence. Learners appear to compare themselves with classmates, whilst teachers use other classes within the institution as a benchmark for their own class's performance. Perceived competence was a poor predictor of actual performance in both the WLC and intervention assessment.

0. The Intervention

The process evaluation was carried out on a pilot of the Welsh language intervention, Automagic, produced by the company *SaySomethingIn*. The intervention uses a micro-lesson format (5-10 minutes input daily) coupled with an inductive approach to learning, with rules and patterns acquired through repetitive language production, rather than explicit instruction. The intervention employs a cyclical model of content development, revisiting existing content whilst gradually integrating new content over time. This cyclical model is used to enhance recombinant language skills, i.e. the learners' ability to rearrange known collocations in different ways to convey meaning, as opposed to the formulaic learning of whole phrases. A video demonstrating the intervention can be found here: https://www.youtube.com/watch?v=p5at_45rVGI

1. The Project Sites and Participants

Seven schools in a diverse, urban area of Cardiff took part in the pilot, with 3 agreeing to take part in the evaluation. Two additional schools (one in Cardiff and one in Caerphilly) were recruited to act as a comparative control group, as baseline data was not available. The pilot schools' pupil cohorts were extremely ethnically and linguistically diverse compared to the Welsh average and reflected higher levels of socio-economic deprivation (Table 1).

Table 1: School demographics and intervention status (data from WG 2023, figures have been rounded to protect anonymity)

School	Status	No. of Learners eligible	Free School Meals	Attendance	EAL	Black, Asian and Minority Ethnic Pupils
Wales Average	**NA**	**NA**	**23.7%**	**91.7%**	**6.7%**	**14.4%**
School 1	Intervention	105	40%	90%	80%	95%
School 2	Intervention	160	45%	90%	80%	90%
School 3	Intervention	100	45%	90%	40%	70%
School 4	Control	35	45%	90%	70%	85%
School 5	Control	30	14%	95%	5%	5%

Across the five schools a total of 417 learners and 16 teachers took part in the research.

2. Methodological Approach

A mixed methods approach was adopted to mitigate for the lack of baseline data and provide insights into the causes of variations in intervention efficacy. Interviews and observations were used to explore and explicate the patterns apparent in the quantitative data. Data outputs included 382 survey responses, 394 oracy assessments, ten in-class observations, 32 learner interviews, and six staff interviews.

Regression modelling was used to explore the relationships between learning outcomes and learner attitudes, and later to inform the interview schedule. Thematic analysis of interview data was then used to identify evidence of possible causative relationships.

3. Limitations of the Research

When considering this research, it is important to be aware of five factors that potentially impact the accuracy and generalisability of the findings:

1. **Limited scope:** the research was conducted on a relatively small scale in a discrete geographical area with a non-randomised sample of schools participating. The results are therefore particular to the specific context of the participating institutions, and may not be indicative of broader regional/national trends. This is particularly pertinent given the atypical demographic make-up of the schools.

2. **Assessment accuracy**: to capture the impact of both the intervention and the continuum learning of participants, the assessment used was broken down into two sections (one assessing continuum content knowledge, and the other assessing intervention content knowledge and recombinant skills). The results from each section cannot therefore be directly compared, but instead are used separately to contrast learning outcomes in each section with other variables of interest (e.g. learner enjoyment). In some instances, the distribution of scores is explored, as this was felt to be indicative of intervention efficacy. The WLC assessment results also do not distinguish between coverage and retention effects i.e. a poor score may be the result of a lack of continuity between the assessment content and the content covered by the learner, or a result of poor learning/retention of content that has been covered.

3. **Lack of baseline data:** it was not possible to collect pre-intervention data of learners' Welsh oracy aptitude. Differences in outcomes between intervention and control groups may be due to pre-existing class/institution variation, rather than the impact of the intervention.

4. **Unknown variables:** whilst the study made efforts to account for confounding variables (e.g. time in Welsh education, curriculum time devoted to Welsh, socio-economic status), there may be unknown variables that have distorted the results.

5. **Hawthorne Effect and Social Desirability Bias:** When conducting research in school settings both these factors may impact on the data collected. The Hawthorne effect is when participants change their behaviour as they are aware they are being observed, whilst Social Desirability Bias can lead to participants reporting what they believe researchers want to hear, rather than their genuine beliefs. Whilst efforts were made to mitigate these factors, it is possible that they have had a distorting effect on the results.

5. Findings

5.1 WLC Assessment Results

Learners generally performed poorly on the continuum section of the assessment (mean score 3.2/40), with over 100 participants scoring zero (Fig. 1). This pattern was consistent across all schools, though intervention schools showed a small but statistically significant uplift of 1.25 marks. However, intervention status accounted for just 3% of the variation in scores overall.
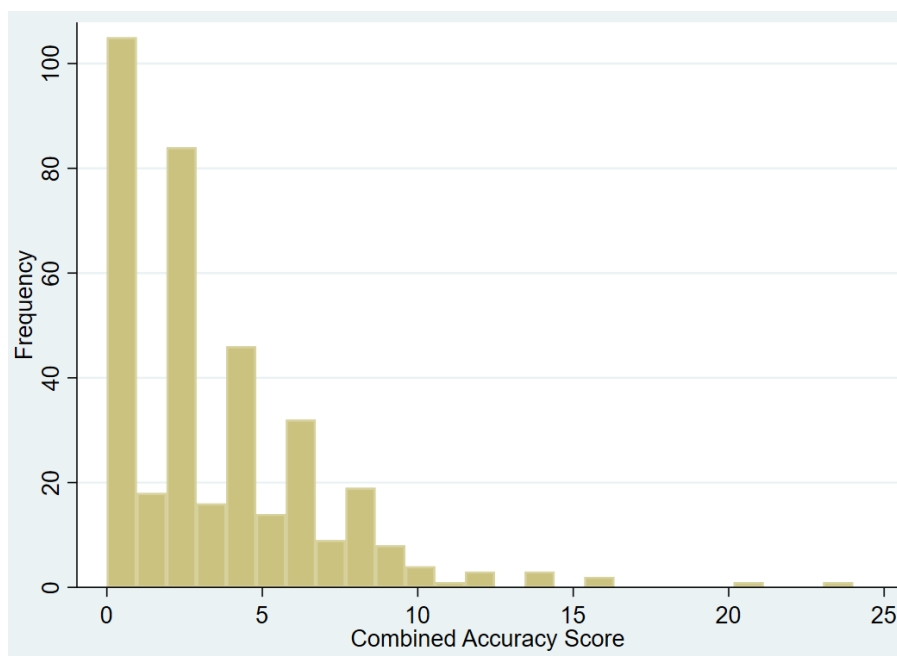

Fig. 1 – Frequency histogram of continuum content accuracy scores

Many learners showed formulaic, rather than recombinant knowledge of WLC content. For example, a student would not know the Welsh for 'it is' in isolation, but would be able to produce a whole phrase containing 'it is', e.g. 'it is sunny today'.

Notably, there was no significant association between perceived academic ability, perceived Welsh language ability, or perceived Welsh importance and continuum content scores. This is surprising as such variables would usually be strong indicators of success (de Fátima Goulão 2014; Lavasani et al. 2010).

5.2 Intervention Assessment Results

In contrast to continuum scores, the intervention assessment scores showed a bimodal distribution (fig 2): whilst the majority learners still demonstrated poor content knowledge, a significant proportion (20%) performed well, scoring 8/10 or higher.
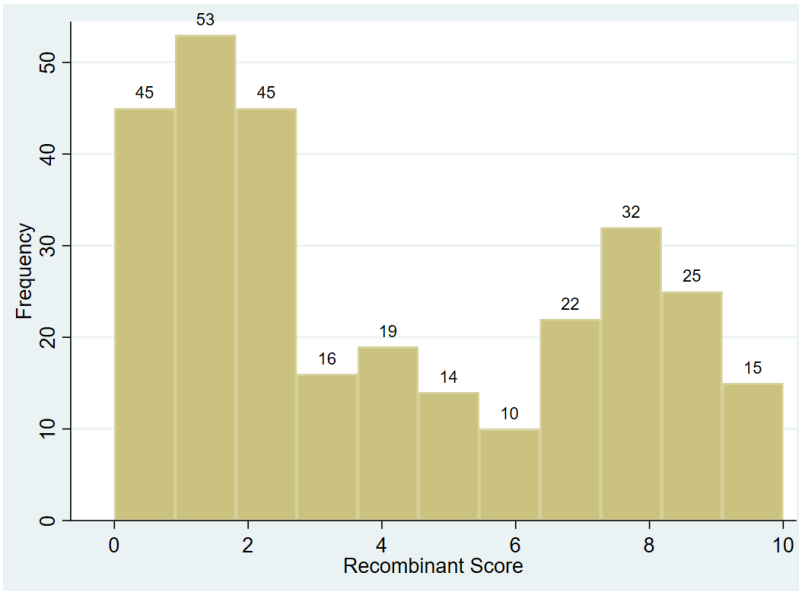

Fig. 2 – Histogram of intervention assessment score frequency excluding control schools

The was no significant association between intervention score and perceived Welsh language or academic proficiency, indicating that learner aptitude was not the defining factor in assessment score. This was supported by staff and learner interviews which indicated that level of engagement was the determiner of success. The larger proportion of learners scoring poorly was ascribed to low levels of engagement, a conjecture supported by lesson observations.

Self-identification with Welsh national identity was positively associated with a 1.1 mark increase in intervention assessment score, concurrent with the interview data that indicated higher levels of integrative motivation in learners identifying as Welsh.

Enjoyment of Welsh was also positively associated with assessment score (0.6 per unit of reported enjoyment), contrasting with the lack of association in continuum scores.

5.3 Learner Self-Efficacy and Sense of Progression

In the absence of formal assessment, learner and teacher perception of competence seems to be largely comparative rather than objective. Interviews indicated that most learners' sense of self-efficacy was formed relative to their peers' performance, whilst staff based their assessment on comparative experience of different classes within the institution. Perceived competence was therefore a poor predictor of actual proficiency.

The effect of the intervention on learner self-efficacy appears to mirror the bimodal learning outcomes. Participants in the intervention group were more likely to consider themselves 'very bad' and 'very good' than those in the control group, who were more likely to cluster around the middle of the scale (Fig. 3).
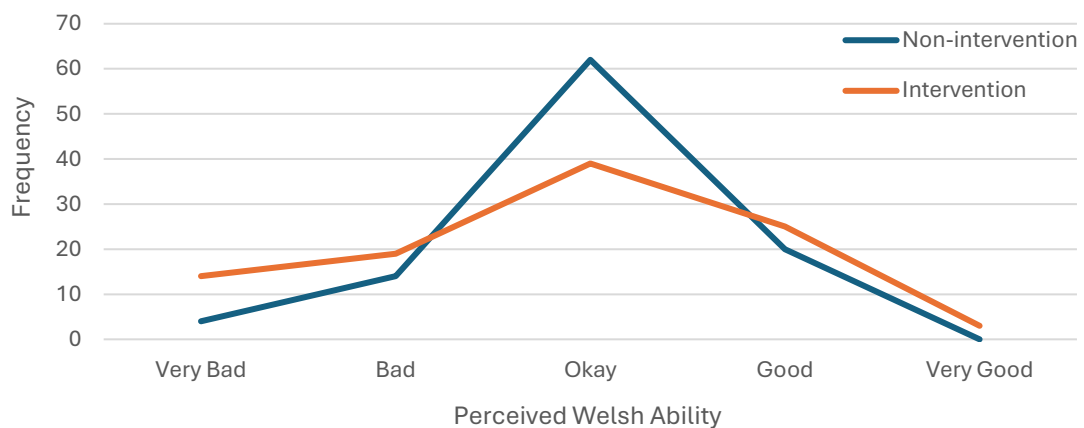

Fig. 3 – Line chart of frequency distribution of perceived Welsh ability by intervention group.

Being part of the intervention group was associated with learners being far more likely to report perceived progress than those in the control group. Although causation cannot be substantiated, and there is likely to be distortion from the Hawthorne effect, the model is consistent with learner interviews, which revealed a high level of perceived efficacy in the intervention, even in cases where learners reported low levels of enjoyment.

Interviews with learners revealed common positive and negative responses to the intervention: whilst many students praised the built-in revision, and short interval 'micro-lesson' learning, others critiqued the intervention for being boring in format, and some complained that the lack of differentiation/individual pacing resulted in them feeling 'left behind'. This was a common critique from teaching staff, who highlighted that the single-stream learning format made re-engaging students who had disengaged or been left-behind (due to periods of absence for example) extremely challenging.

The high levels of non-Welsh identity in the schools appears to result in additional challenges in learner motivation, with those with non-Welsh cultural and linguistic backgrounds being less likely to draw on integrative motivation. Teachers sought to compensate for this by emphasising the instrumental utility of the language (e.g. job prospects, qualifications), and enhancing lesson enjoyment through gamification to generate intrinsic motivation.

6. Recommendations

Recommendation 1: It appears that the difference in score distributions between WLC and intervention assessments may be a product of poor content retention/acquisition. The research suggests this may stems from a combination of poor levels of engagement/motivation, and insufficient revision/integration of previous content. The recursive model used by the Automagic system appears to go some way to addressing this. Such an approach could be adopted into WLC lessons through the enhanced use of the daily Helpwr Heddiw sessions. This would reflect the learner preference for the microlesson format, and potentially address the challenge of learner disengagement resulting from a lack of perceived progress (manifesting as poor self-efficacy). Staff would benefit from training in planning such an approach.

Recommendation 2: The WLC patterns learners have acquired often appear to be formulaic, leading to limited communicative flexibility. The Automagic intervention's focus on recombinant 'chunking' of language requires the learners to create new patterns from existing collocations, thereby improving recombinant skills, and transferring patterns into procedural memory by mimicking the meaning-focus retrieval of authentic usage. Upskilling staff to utilise such an approach with the continuum content would potentially lead to higher levels of content retention and long-term communicative efficacy.

Recommendation 3: The lack of assessment tools appears to have resulted in a highly contextualised understanding of proficiency. The development of a standardised assessment tool would provide teachers and schools with a better understanding of the development of Welsh language skills within their class/institution, enabling them to direct teaching more effectively towards gaps in learner proficiency.

Recommendation 4: Whilst the intervention was praised by staff for its ease of use, the lack of differentiation or scaffolding was highlighted as exacerbating learner disengagement. Adaptations allowing for individual self/teacher differentiation, scaffolding for lower achieving learners, or the integration of an Adaptive Learning Technology (Bradac & Walek 2017, Heil et al. 2016) may help mitigate the polarising effects observed in the learning outcomes and learner self-efficacy.

Recommendation 5: Many learners were critical of the intervention's format as being 'boring' and not varied enough to hold their interest. Learner engagement may be enhanced through the integration of game-like features, or competitive elements.

References

Bradac, V. and Walek, B. 2017. A comprehensive adaptive system for e-learning of foreign languages. *Expert systems with applications*, *90*, pp.414-426.

de Fátima Goulão, M. 2014. The Relationship between Self-Efficacy and Academic Achievement in Adults' Learners. *Athen's journal of Education*, *1*(3), pp.237-246.

Heil, C.R., Wu, J.S., Lee, J.J. and Schmidt, T. 2016. A review of mobile language learning applications: Trends, challenges, and opportunities. *The EuroCALL Review*, *24*(2), pp.32-50.

Lavasani, M.G., Malahmadi, E. and Amani, J. 2010. The role of self-efficacy, task value, and achievement goals in predicting learning approaches and mathematics achievement. *Procedia-Social and Behavioral Sciences*, *5*, pp.942-947.

Welsh Government. 2023a. *My Local School: Sharing Information about Local School*. Available at: https://mylocalschool.gov.wales/School/ [Accessed: 3 August 2024]