Using text-mining methods and a genealogical discourse analysis, our comparative study aims to explore the views of Waldorf and Montessori teacher trainers. The data was collected through structured interviews. The sample consists of eight respondents involved in Montessori teacher training and eight respondents involved in Waldorf teacher training ($N_{\text{total}} = 16$). Our research objective is to identify the lexical and narratological features, similarities and differences in the monologues of expert teachers with the aim of their interpretive evaluation. The literature review provides a brief summary of recent research on didactics and effectiveness, as well as some critiques of Montessori and Waldorf pedagogy.

1. Literature review

1.1. Montessori education
Montessori teachers work with tools and prepared environments (Jurčík, 2023). Maria Montessori interpreted the child’s play activity as a workflow, so that in Montessori pedagogy there is no distance between ‘work’ and ‘play’ in the usual sense, the two are conflated in the methodology, which can be seen in two ways. On the one hand, Lillard (2021) sees a limitation to the expansion of Montessori pedagogy in the fact that it may be perceived as too work- or play-like, and on the other hand, Montessori students have more positive attitudes towards learning than mainstream students, and while traditional school graduates were most likely to enjoy being with friends, Montessori students were most likely to enjoy learning in school (Snyder et al., 2023). Montessori educators consider mathematics and literacy to be key elements of the programme (Beach, 2023), so Montessori pedagogy contributes most to mathematical and literacy skills (Ansari & Winsler, 2022; Gasco-Txabarri & Zuazagoitia, 2022), while Montessori methods do not limit the directions of a child’s development and promote emotional development (Lapon, 2020). Montessori education can be applied to children with a range of difficulties (Pérez-Pérez et al., 2021), and Montessori teachers incorporate research developments into their work with SEN students (AuCoin & Berger, 2021). Montessori pedagogy places great emphasis on music.
in the education of children, but music education does not receive as much attention as other areas such as language, mathematics and science.¹

Even though empirical studies of the performance of Montessori schools on standardised tests have been inconsistent (L. Snyder et al., 2022), there is evidence to suggest that Montessori pedagogy, properly implemented, is effective (Courtier et al., 2021; Randolph et al., 2023; Tiryaki et al., 2021). The impact of the implementation level of Montessori pedagogy on the validity of outcome measures is examined by Basargekar & Lillard (2021). Lillard (2019) argues that the generally positive results of outcome tests are one of the external reasons for the persistence of the Montessori concept. The results of some studies are limited by the study by Gentaz & Richard (2022), in which the authors show that the only three RCT studies published so far do not include a fully active control group using alternative pedagogy and do not examine the quality of implementation of Montessori pedagogy. The results of some studies are limited by the study by Gentaz & Richard (2022), in which the authors show that only three RCT studies published to date do not include a fully active control group using alternative pedagogy and do not examine the quality of implementation of Montessori pedagogy.

1.2. Waldorf education

There are a number of criticisms of Waldorf Education in the literature, some of which will be discussed below. A typical set of criticisms are publications that question the authenticity of Steiner’s ideas. Some authors suggest that it is a reworking of the ideas of others: Kutschera (2016) suggests that the use of the word ‘biodynamic’ in anthroposophy is a misuse of a concept from Ernst Haeckel’s pioneering work of 1866, a distortion of the original meaning. French (2021, pp. 20–27) points out the similarities between Steiner and Max Weber in their critiques of technological progress, and between Steiner and Nietzsche in their ideas on ethical individualism (2021, pp. 69–80). Zumdick (2019) shows significant connections between the foundations of Waldorf education and Schiller’s aesthetic principles of education, and Rijke (2019) between the foundations of Waldorf education and Vygotsky’s developmental psychology. According to Beyerle & Beyerle (2018), the idea that the purpose and destiny of humanity is the realisation of the ‘I’ appeared in Jung before Steiner, and Hindes (Hindes, 2021) and Staudenmaier (2008) suggest that theories of root species, lost continents, national karma and the destiny and evolution of the human races appeared in Theosophy. The theory of the destiny and evolution of the human races is the subject of a separate, larger set of criticisms. They argue that Steiner

¹ Montessori, following Dalcroze, believed that the performance and reproduction of music could only take place after the child had understood and experienced music through movement Laure and Habe (2023).
promoted racist ideas, and even that extreme right-wing elements have always been present within anthroposophy (Staudenmaier, 2014, pp. 319–327). This criticism is so typical that it appears not only in German and in professional or social discourse, but also in studies on other subjects (Majerus, 2022; Muhie, 2022).² According to Berner (2022), eurythmy was a child of the flourishing Körperkultur of the time, Émile Jaques-Dalcroze held his first eurhythmics class in the autumn of 1903 and then introduced his eurhythmics’ method in several places in Switzerland and Western Europe between 1903 and 1910 (Montoya Rubio, 2023), eight years before Steiner (eurythmy': 1911-1912). Before Steiner, Dalcroze also declared that the primary musical instrument was the human body (Daly, 2022).

Without giving further examples, many scholars argue or suggest that Rudolf Steiner simply applied the ideas of his time to anthroposophy by rhetorically and conceptually restructuring them. Other critics (Wilson, 2014, 2022) argue that Waldorf pedagogy is only nominally a pedagogy of freedom, in reality it is defined by routines and rules. However, Stehlik (Stehlik, 2019b) argues that Waldorf teachers are explicitly expected to be creative in their curriculum and methods, which is confirmed by everyday experience. Despite the criticism of the anthroposophical background, the results in the applied fields of anthroposophy are positive. For reasons of space³, we will concentrate here on two examples: anthroposophic medicine and biodynamic agriculture.

Anthroposophic medicine (AM) is based on Rudolf Steiner’s anthroposophy and complements conventional medicine. The scientific status of AM is adequate (Baars et al., 2022), and indeed the first guidelines and standards for quality assurance have recently been published (World Health Organization, 2023). Anthroposophic medicines are produced in accordance with good manufacturing practice and national drug regulations, and the limited evidence available suggests a clinical benefit (Baars et al., 2022). Although the positive effects are not yet clear in all areas, anthroposophic medicine is effective (Ploesser & Martin, 2023).

Biodynamic agriculture, which is based on collective values and aims to be a rational alternative to the industrial model (Wright, 2022), is a cooperative approach where anthroposophy meets small communities (Ganany-Dagan, 2022). Biodynamic farming is now practised by more than 5,500 farmers worldwide, with 41.8% of the world’s total area in Germany (Banerjee & Saha, 2023). Organic farming is practised on about 10% of the total agricultural area in Denmark (Aare et al., 2021). Although costs are higher and yields

---

² The Dutch government commission that had been set up to investigate this accusation found that in 0.05% of Rudolf Steiner’s 89,000 pages of published works there were a total of 16 sentences that, as isolated statements, violated the modern, well-developed and highly sensitive Dutch discrimination law [Stehlik (2019a)].

³ There are many other examples. For example, a health-conscious behaviour of Waldorf pupils is more developed and the proportion of overweight children is lower than that of mainstream pupils [Blaychfeld Magnazi et al. (2023)]. Vitale & Coccia (2023) found a similar phenomenon among Montessori students.
are lower compared to conventional solutions (Robusti et al., 2020), a literature review of 86 studies shows that biodynamic farming has proven positive effects on agro-ecosystems and food quality (Brock et al., 2019). Accordingly, some authors (Rigolot & Quantin, 2022) aim to contribute to dispelling common misconceptions about biodynamic farming and to establishing a more constructive dialogue with the wider scientific community.

2. Method

2.1. Study objectives, measuring instrument, sample
The data was collected through structured interviews. The 24 questions correspond to the curricular themes revealed by the literature review mentioned in the introduction. The questionnaire, the information sheet, the full text of the interviews and the data files for analysis are available at https://doi.org/10.7910/DVN/S96QV9. The sample consists of eight respondents involved in Montessori teacher training and eight respondents involved in Waldorf teacher training ($N_{total} = 16$). The respondents were selected using a snowball and criterion-based qualitative sampling strategy. Criteria included having state and school model qualifications, relevant teaching experience, involvement in teacher training and activity in the organisation. The snowball sampling strategy focused on respondents’ awareness and significant theoretical knowledge. The sample covers a significant proportion of the population (60-90%). Respondents gave prior informed consent and the study met ethical requirements. Our research objective is to identify the lexical and narratological features, similarities and differences in the monologues of expert teachers with the aim of their interpretive evaluation.

2.2. Lexical characteristics of the responses
In the first step of the analysis, the descriptive statistics, lexical diversity and similarity of the corpora were examined in a comparison of the two school models. The validity of the conclusions drawn from the differences in lexical characteristics was increased by the fact that the interviewees were allowed to give their answers within the framework of the structured interview with unlimited time, i.e. the length and internal diversity of the answers depended on the interviewees. As four sub-questions were addressed exclusively to Waldorf teachers, questions 13, 14, 23 and 24 their answers were deleted from the Waldorf corpus for analysis. The accuracy of the transcriptions of the audio recordings was checked by two persons independent of each other. Prior to the analysis, the data was cleaned: noise (punctuation, special characters) was filtered out of the corpus, words were lowercased, punctuation was removed, the Hungarian Stop Dictionary (HunMineR) was
used to remove words from the disallow list, and finally lemmatisation was performed. The descriptive statistics of the corpora showed the length of each response alone and as an average of the Montessori and Waldorf subsamples. Since the responses in the Waldorf subsample were longer on average, we used the Maas and MSTTR indices, which have been shown to be more robust than TTR (Fergadiotis et al., 2015; McCarthy & Jarvis, 2010; Torruella & Capsada, 2013), to calculate lexical diversity. To ensure the validity of the lexical diversity claims, the results of both computational procedures were considered together. After detecting the differences, we looked for the extent to which a respondent gives lexically similar answers to different questions within the whole interview. We also looked for the same in the average of the two subsamples. As it is not sensitive to the length of the text, we used a cosine method of measurement. Although the cosine values obtained are not easy to interpret on their own, they are fully suitable for comparisons within and between the Waldorf and Montessori subsamples. Similarity was examined not only within each response but also between responses.

However, the lexical features of a freely performed linguistic production do not only depend on the depth and breadth of knowledge, but also on the constant and current speaking attitude of the respondent (e.g. tendency to conciseness or colouring with stories), age and education, and on the ability to deliver a monologue in general. Furthermore, the measurable impact of a diverse and even lengthy response may actually be less than the professional value of a less lexically complex and even short response. In other words, the professional value of an answer cannot be determined by examining word repetition. The general research question cannot be validly answered by examining lexemes alone, but it does contribute to the final results. In order to draw valid conclusions, we also found it necessary to examine the narratological characteristics of the responses.

2.3. Narratological characteristics of the responses
A monologue is not a discourse with a single participant, a monologue requires a recipient, so there is a social stake in the speech, and an expert respondent is more or less likely to claim legitimate ownership of his or her subject. However, a story told to others not only describes reality but also creates it. Since the genealogical approach does not try to tell us who we really are, but how we conceive of ourselves as subjects, how we have created and transformed ourselves into political beings, we used a genealogical nature of discourse analysis (Anaïs, 2013) in the second step of our analysis. In the process of examining the discourse as a whole, we looked for and explored in an interpretative way how a need for knowledge, reality and claims to truth interact. We compared key meanings and chains of meaning in the texts, looking for points of rupture in the discourse. We looked at whether the concepts used by the interviewees fit with the theory or whether they built their concepts
out of the social space. We also saw the scientific stakes and political significance of the methodological momentum by examining: is there a profession behind the role?

3. Results

3.1. Lexical features of the answers

According to the descriptive statistics of the corpora, the response length of the least talkative Montessori expert is 24893 characters, and the response length of the least talkative Waldorf expert is 47045 characters. The most talkative Montessori expert has an answer length of 56970 characters, the most talkative Waldorf expert has an answer length of 114980 characters. On average the Waldorf respondents answered the same questions 78% longer [Table 1].

<table>
<thead>
<tr>
<th>Montessori respondents</th>
<th>Response length (characters)</th>
<th>Waldorf respondents</th>
<th>Response length (characters)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>41566</td>
<td>W1</td>
<td>49865</td>
</tr>
<tr>
<td>M2</td>
<td>37015</td>
<td>W2</td>
<td>65793</td>
</tr>
<tr>
<td>M3</td>
<td>42708</td>
<td>W3</td>
<td>47045</td>
</tr>
<tr>
<td>M4</td>
<td>56970</td>
<td>W4</td>
<td>79847</td>
</tr>
<tr>
<td>M5</td>
<td>33289</td>
<td>W5</td>
<td>65978</td>
</tr>
<tr>
<td>M6</td>
<td>35017</td>
<td>W6</td>
<td>114980</td>
</tr>
<tr>
<td>M7</td>
<td>24893</td>
<td>W7</td>
<td>50318</td>
</tr>
<tr>
<td>M8</td>
<td>32795</td>
<td>W8</td>
<td>67696</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td><strong>38031.62</strong></td>
<td></td>
<td><strong>67690.25</strong></td>
</tr>
</tbody>
</table>

The lexical diversity calculated by two different methods, robust to the length of the corpora, gave similar results: both groups are homogeneous, with no significant difference between the means of the Waldorf and Montessori subsamples [Figure 1].
As lexical diversity was not examined in a time series but in a cross-sectional arrangement, no firm conclusions about the research question can be drawn from individual performance and the difference in means between groups. However, it is significant that lexical diversity is dependent on the age and level of education of the respondent and, despite the list of inclusion criteria, there were differences between respondents in these two aspects. Nevertheless, the respondents within the groups and the average of the two groups are similar in terms of word repetitions (skipped sentences, repeated statements, omissions and insertions, etc.). The results of the lexical diversity analysis confirm the validity of the comparative analysis.

Following lexical diversity, the „internal” similarity of the answers was examined to see if the diversity of the answers at the level of the individual and on average between the two groups was related to the diversity of the questions. Examining the internal similarity of the responses contributes to answering the research question already in terms of content. The results show that the scores within the two groups are similar but not homogeneous. On average, the Waldorf subsample gave more lexically similar answers to the different questions than the Montessori subsample [Figure 2].
On average, therefore, the Waldorf sample answered more of the different questions with the same set of vocabulary. The result suggests that the Waldorf subsample has a smaller diameter of the 'theoretical bubble', which can be evaluated in a number of ways. It is possible that the Waldorf respondents are more self-repeating at the lexical level because their knowledge is more monolithic than that of the Montessori respondents. It is also possible, however, that Waldorf respondents are more stable and consistent in their theoretical framework, which allows them to answer different phenomena (questions) in a more consistent and less diverse way. A firm decision between the two options is not possible by examining the internal lexical similarity of the responses.

A more informed decision can be made based on a two-step test of 'external similarity'. In the first step, we used cosine similarity to find the average similarity of the answers given by respondents in the two groups compared to the rest of their own group and in comparing the groups [Figure 3]. The results show that the Waldorf respondents’ responses were more similar to those within their own pedagogical subgroup than the Montessori respondents. Therefore, of the two possibilities indicated by internal similarity, the narrative that Waldorf respondents are more stable and consistent in their connection to their theoretical framework is the more likely.
In a second step, we used hierarchical clustering of the responses using Euclidean distance to see how far the two groups were separated by the lexical features of the responses. The interpretation of the dendogram obtained from this procedure allows us to show how members of different groups are clustered according to their word use, what overlap exists between them, whether group members (e.g. some Waldorf experts) can also be grouped to another group (in this example the Montessori experts) based on their word use [Figure 4].

On the basis of the nearest neighbour’s classification, the Montessori and Waldorf models are well separated. In the Waldorf subsample, a pair (W1, W3) is joined by respondent W7 to form a cluster of five pairs of respondents W5 and W8. This cluster is joined by respondents W4 and W2. In the Montessori model, three pairs of three (M2
and M1, M8 and M6, M5 and M7) are joined by M3 and then M4 in a cluster of six. It is noteworthy that the two sub-samples are joined by a Waldorf expert respondent (W6). The results therefore suggest that the separation of the two groups is almost perfect.

To summarise the analysis of the lexical characteristics of the answers, the monologues were significantly longer in the Waldorf sub-sample. The lexical diversity between the groups, independent of the length of the answers, was the same in the Waldorf and Montessori sub-samples, which confirms the validity of the comparative analysis. For the Waldorf subsample, the mean internal similarity of responses was slightly higher, the reason for which was determined in a second procedure. The results suggest that Waldorf respondents are more stable and consistent in adhering to their theoretical framework, which allows them to answer different phenomena (questions) in a more consistent and less diverse way. This result may be related to the differences in the available training opportunities of the two models in Hungary. Finally, the classification of the answers showed that the two groups are not lexically homogeneous and although one Waldorf expert can be connected to both groups, the experts of the models are divided into two groups without overlapping.

3.2. Narratological characteristics of the responses
When asked how original the Waldorf/Montessori theory and methodology is, one Waldorf teacher (W3) answered in the negative:

I don’t think so. It tries to touch, respect and cultivate thousands of years of mature knowledge and very few people are able to see through that arc and cultivate it to the standard the founder had in mind.

The need for lifelong development and professional commitment characterise both pedagogies. A Montessori respondent (M4) commented:

I had a professor, now 86 years old, who said it takes 7-10 years to be a good Montessori teacher, but there were people who said 20. So it’s a long journey, you go through a transformation and I think you are not a Montessori teacher, you are a Montessori person or a Montessori adult. You are not applying Montessori, you are living Montessori.

The Montessori respondents had a similar view of the relationship of their pedagogy to art and movement and the directions and possibilities for development. Although its effectiveness and values were seen as unquestionable, none of the interviewees called for its
implementation in mainstream education, but at most recommended it. The social status and role of Waldorf and Montessori education was seen by the interviewees in a nuanced, limited and reflective way. The question of Rudolf Steiner’s anti-Semitism was received with genuine dismay by the Waldorf respondents and was then clearly rejected on the basis of the interviewees’ reading experiences and their own and others’ daily practice. Only one interviewee (W8) was aware of such criticism. As the literature had not yet been published at the time of the data collection and the research was based on themes from the existing literature, the Montessori respondents were not asked about the criticism accusing Maria Montessori of anti-Semitism. The Montessori interviewees clearly expected much more organisational effort from the organisers and leaders of the movement in Hungary in order to better publicise the pedagogy and foster community relations. In both models, the aim is to create state-recognised diploma courses and to carry out research, development and professional accreditation in higher education.

4. Discussion and conclusion

Consistent with theory (Tsorantidou et al., 2021), Waldorf respondents clearly stated that it is the low-tech technological tools that are typically used in Waldorf pedagogy. According to Marshall (2017), Montessori teachers often feel uncertain about their ability to innovate their practice while maintaining the principles of Montessori pedagogy. In contrast, in the present study, there was no perceived uncertainty in the responses of Montessori practitioners, with respondents clearly accepting the incorporation of modern technological tools into their pedagogical practice (Ahmed Sayed Ali et al., 2021), with age restrictions and if those tools are consistent with Montessori principles. Given the recent increase in proposals to ban mobile devices from classrooms, in contrast to the uncontroversial support for mobile devices in education (Selwyn & Aagaard, 2021), the results suggest that neither Waldorf nor Montessori pedagogy can be accused of stagnating and failing to meet societal expectations. Although Waldorf methods implemented within the framework of pluralist pedagogy have shown positive results (Shank, 2016), there were Waldorf respondents who explicitly disagreed that Waldorf pedagogy should definitely be adapted to pluralist pedagogy (Wiehl, 2021). However, several Waldorf educators referred to the concept of salutogenesis (David, 2019), a more positive attitude of students towards school and learning. Some also referred to possibly poorer performance (Salchegger et al., 2021), but nobody referred to findings that may contradict the idea about students’ more positive attitudes towards learning (van Schie & Vedder, 2023). Despite the lack of reflection in the literature on achievement, teachers in both models were self-reflective, in line with their leadership
role and associated professional expectations (Dames, 2019; Damore & Rieckhoff, 2021). In relation to leadership and professional expectations, Montessori respondents also confirmed that they navigate a system on a daily basis that is often inconsistent with their pedagogy, i.e. while effectiveness depends on the degree of implementation of the original form of pedagogy (Gerker, 2023), circumstances more often force them to deviate from high quality implementation.

Despite the fact that the average educational level of the Montessori respondents was lower than that of the Waldorf respondents, the lexical diversity of the answers of the two groups was the same. An analysis of the similarity of the responses showed that the Waldorf respondents were more grounded and consistent in their theoretical framework, which may be related to the differences in the educational provision of the two models in Hungary. The classification of the responses showed that the two groups are not lexically homogeneous, and although one Waldorf expert carries both groups’ characteristics, the experts of the models can be divided into two groups without overlap.

Three horizons of meaning for Waldorf and Montessori education have been identified on the basis of the texts and the ways in which they are expressed. The first is the social narrative of school models, the everyday stories told about the pedagogical concept, shaped mainly by the media, opinion formers and school users. It is this, rather than scientific evidence and refutations, that contributes most to lay acceptance and, above all, determines the number of applicants to institutions. Second, the professional expertise of the teachers, the quality of their daily practice and of their problem-solving strategies, which has professional and educational policy implications beyond lay reporting. Thirdly, theory in its pure form, which has an important role to play if it is nurtured, tested and passed on by expert teachers who know and practise the profession.

References


