Exploring music and sound in multimodal literacy: A systematic review and its implications for music education

José Juan Roa-Trejo¹
Dpto. de Comunicación y Educación, Universidad Loyola Andalucía, Sevilla (España)

Alejandra Pacheco-Costa²
Dpto. de Educación Artística, Universidad de Sevilla, Sevilla (España)

Francisco Cuadrado³
Dpto. de Comunicación y Educación, Universidad Loyola Andalucía, Sevilla (España)

doi: 10.7203/LEEME.52.27500
Reception: 04-10-2023 Revision: 05-10-2023 Acceptance: 10-11-2023

Abstract
Multimodal literacy involves using multiple communication modes in text comprehension and production. This review examines sound and music in research on multimodal literacy, analyzing 48 pieces from Web of Science and Scopus databases. It identifies common methods and educational aspects in the texts and explores how multimodal literacy is applied in academic settings, focusing on the implications that it may have for music education. The study highlights the versatility of the aural mode, encompassing sound effects, vocalizations, music, and more. However, some studies merely use sound as "background music" without considering its communicative role. This research contributes to understanding sound's implications and its potential in engaging students in a multimodal society.

Key words: Multiple Literacies; Sound; Music; Education.

Resumen
La alfabetización multimodal implica el uso de múltiples modos de comunicación en la comprensión y producción de textos. Esta revisión examina el sonido y la música en la investigación sobre alfabetización multimodal, analizando 48 trabajos de las bases de datos Web of Science y Scopus. Identifica métodos y aspectos educativos comunes en los textos y explora cómo se aplica la alfabetización multimodal en entornos educativos, poniendo el foco en las implicaciones que esto pudiera tener para la educación musical. El estudio destaca la versatilidad del modo auditivo, que abarca efectos sonoros, vocalizaciones, música y otros. Sin embargo, algunos estudios limitan el uso del sonido a la "música de fondo", sin profundizar demasiado en su función comunicativa. Esta investigación contribuye a comprender las implicaciones del sonido y su potencial para implicar a los alumnos en una sociedad multimodal.

Palabras claves: alfabetización; sonido; música; educación.

¹Ayudante de Investigación, Facultad de Psicología y Educación, https://orcid.org/0000-0002-4207-1160
²Profesor Titular de Universidad, Facultad de Ciencias de la Educación, https://orcid.org/0000-0001-6397-4708
³Profesor Titular de Universidad, Facultad de Comunicación y Artes, https://orcid.org/0000-0003-2307-3846
*Contacto y correspondencia: José Juan Roa-Trejo, Departamento de Comunicación y Educación, Facultad de Psicología y Educación, Universidad Loyola Andalucía, jjroa@uloyola.es, Avda. de las Universidades, 2, 41704 Dos Hermanas, Sevilla. Spain.
1. Introduction

We are in a time of constant change, especially in the field of education. The shift from pencil and paper to screens, the emergence of Recurrent Neural Networks for text generation (such as ChatGPT), or the proliferation of media such as audiobooks, social networks, video game chats and streaming platforms or forums, are events that lead us to rethink what types of language should be taught in educational institutions (Domingo, et al., 2015; Marsh, 2004). The European Literacy Policy Network (ELINET) already recognizes that literacy should include digital media drawing on multimodal skills and knowledge, such as visual and auditory information production (Valtin, et al., 2016). In the last 40 years, studies have emerged that define literacy as a socio-cultural practice, conceptualizing literacy as embedded in society and built upon vernacular practices, whether analogue or digital (Street, 1985; 2011). Thus, there is a degree of agreement within the research community that literacy research does not only involve identifying the psychological processes that take place within the individual which underpin reading and writing (Bloome, & Green, 2015) but also requires recognition of all those processes in which decoding occurs and through which meaning is created (Gillen, & Hall, 2013).

Changes in ICT and the emergence of new literacy theories pose significant challenges in translating these concepts into the classroom. According to Bazalgette and Buckingham (2013), terms such as “multimedia”, “digital” or “multimodal” literacy are used as “hooks” for advertising purposes, without having real applications in practice. Furthermore, they propose that teachers feel alienated from students’ textual practices owing to changes in communication technologies, which hinders the teaching-learning process. There are also difficulties when it comes to materializing multimodal literacy approaches in the classroom, as they are built on varied definitions of literacy (Nash, 2018). One challenge for teachers is assessing how successful their approaches to multimodal literacy are since, according to Anderson and Kachorsky (2019), teachers often use “high-stakes forms of summative assessments” (p.313) which are not aligned with ideological models of literacy. Moreover, designing and implementing educational strategies that address different forms of communication can lead teachers to a situation of uncertainty, so a certain level of training and specialisation is necessary (Kuby, et al., 2015).

These difficulties are exacerbated when working with code systems that include sound. Owing to the immaterial and abstract properties of sound and music, these “are not considered as powerful conveyors of meaning and are not used to teach communication skills” (Rowsell, 2013, p.148). According to Shanahan (2012), sound is generally used as a “decoration” within the literacy framework in the classroom, without delving into its communicative capabilities. Shanahan argues that, despite its recognition as a mode of communication, there is a lack of research into the analysis of sound and its implications for a message. Technological changes and the permeability of the boundaries between sound, music and other forms of communication make it difficult to use and analyze the semiotics of the aural (van Leeuwen, 1999). As we are constantly surrounded by sounds, it is essential to know the aural contexts that surround pupils, as well as the codes that are used in them and how they influence the meaning-making processes that take place in childhood (Petchauer, 2020).

Other authors have performed literature reviews about multimodality at school settings. Aforementioned Anderson and Kachorsky (2019) carried out accurate research about the different methods used to assess multimodal proposals. On the other hand, Lim, Toh and Nguyen (2022)
examined the existing literature about multimodality at the English teaching context. With the intention of shedding light on the possibilities of sound and music within the new conceptualizations of literacy, as they haven’t been explored by other authors, the present systematic review was undertaken. To this end, we set out the following research objectives:

- Describe the main characteristics of the studies that make up the scientific corpus of multimodal literacy, both in the academic and educational fields.
- Review and evaluate the semiotic implications of sound and music in the multimodal paradigm addressed by these studies.
- Collate the findings and possible improvements from the sample studied for the purpose of defining a possible line of evolution for the field of multimodal literacy studies that include sound elements.

1.2. Multimodal Literacy and Sound

Literacy is an active process of meaning-making through which the world is recognized (Street, 2011). Thus, literacy is influenced by the way new technologies conceive literacy practices and is therefore multimodal (Rowsell, & Walsh, 2011). The multimodal property of literacy is that communicators create meaning through different types of modes (Flewitt, 2013). Modes are systems of codes that achieve an effect in communicational processes (image, gesture, sound, movement, etc.). The modes used in communication directly affect the message, even interacting with each other in different ways (Kress, 2010). Communication then “occurs through the combination of juxtaposed sign systems to create a more powerful effect” (Crafton, et al., 2009, p.34). Thus, modes are selected and redesigned to shape the meaning conveyed in any social and cultural medium (Tomlinson, 2015). According to Kress (2010), a mode is a semiotic resource used to make meaning. A mode is also defined by social needs of communication. For example, a graphic designer will choose a particular typeface depending on their needs and knowledge of what a particular community understands in relation to conventional practices about typefaces (Bezemer, & Kress, 2008).

Multimodal literacy opens the door to children's critical awareness, and numerous studies use the multimodal perspective to incorporate learners' contexts, recognizing the social meaning of literacy practice (Simpson, & Walsh, 2015). At times, this multimodal perspective carries an implicit artistic view of the communicative act that engages teachers and learners in deconstructing traditional discourses and assists them in creating new codes (Binder, & Kostopoulos, 2011). Pahl (2019) proposes that through the inclusion and study of artistic elements in communication, elements that bring the language of learners closer to educational system are recognised. The development of multimodal literacy skills is crucial in collectives that construct their identity through oral language and elements such as styles of music, forms of dress, or visual artistic manifestations such as graffiti (Mills, & Unsworth, 2018). It is necessary to recognize that learners naturally make complex use of different modes in order to include the communicative context of learners in the literacy processes that take place within the classroom (Harrop-Allin, 2017). One of the strategies that can assist learners in reflecting on these resources is “remediation”. Remediation is the adaptation of a message from one mode to another, recognizing how this change affects meaning (Alexander et al., 2016). This remediation offers opportunities to build connections between different domains that make different literacies holistic and meaningful (DePalma, 2015). Morera et al., (2020) proposed an interesting example of
remediation when moved from conventional choral music towards polyphonic sign language in pursuit of inclusion.

Working with multimodal elements makes it possible to describe, explain and explore how semiotic resources can offer more options and tools for the production and interpretation of meaningful actions (Kress, 2010). Van Leeuwen (1999) worked on describing the semiotic possibilities of sound (musical and non-musical), dividing them into six domains:

1. Sound perspective: involves the spatial relationship between sender and receiver and distinguishes positioning and direction, allowing the creation of a soundscape, as well as immersion in it. It distinguishes the intimate (close, like a whisper) from the informal (medium, relaxed) and the formal (louder, a projected message to the public).
2. Time and rhythm: time can be free or adjusted to a measure, thus defining rhythm. We can find simple rhythms or polyrhythms in both musical and textual discourse. Metrical patterns, such as beats, have a meaning, but so does the lack of organised structures.
3. Interaction of voices: allows the sequencing of sound and the set of semiotic meanings derived from each of them (hierarchy, segregation, complementarity, heterogeneity, etc.).
4. Melody: the pitch and its alterations communicate meanings, but also continuity, range, articulation, etc.
5. Vocal quality and timbre: timbre is multidimensional. Each sound is a mixture of different characteristics. A voice is high-pitched or low-pitched, or soft or intense. To describe timbre, it is necessary to go beyond adjectives (metallic, shrill, hollow) and consider the physical sound, i.e., as it is materially produced.
6. Modality: lies in sound's capacity for representation. The term modality refers to the degree of "truth" assigned to a sound event. This means that some sound events are understood as "less true" than others depending on the "modal judgement" of the listener, based on its aural representations of different ideas. Modality is based on the criterion of similarity: the more the sound resembles how it should sound if the represented event were present, the higher its modality.

The categorization of sound's communicative potential holds significant value in the evaluation of the semiotic roles played by aural semiotic resources in the context of multimodal literacy practices. It might be possible to learn more about the implications of sound in the relationships established between subject, signifiers, and meaning (Nowak, & Bennett, 2014). In fact, some authors go further in the conceptualization of sound and its connection with other modes or disciplines. Powell and Somerville (2020), for example, analyze children’s sound experience from an ontological approach, considering sound and music inextricably linked with movement and place. Conceptualizing sound as a sensory experience leads to challenge the relation signifier-signified relation and to recognize the complexity of musical interactions, specifically in early ages of childhood (Hackett, & Somerville, 2017).
2. Method

In order to explore the extent to which and in what ways sound is considered by authors within multimodal literacy approaches, a systematic review was conducted. We adopted the PRISMA approach, given its usefulness and the low risk of bias (Shamseer, et al., 2015; Page, et al., 2021). PRISMA is an approach to systematic review that enhances the rigor and reliability of evidence-based research by establishing four steps in the review process: Identification, Screening, Eligibility and Inclusion. However, slight modifications were made to some of the tools it proposes, with the intention of fine tuning the search and reducing bias. For example, the method of inclusion of publications has been adapted to the sample, as we did not use a checklist but a peer review system through which three co-authors validated the inclusion or exclusion of all sample elements.

2.1. Review process

Firstly, a search was conducted of the Scopus and Web of Science (WoS) databases. English descriptors were used to give the research an international scope and the search was confined to the period 2012 to 2023 (inclusive) to ensure that the findings were up to date. The following search terms were used in the title, abstract and keywords: "multimodal", "literacy" and "sound" or "music" as well as related terms, such as plurals or multimodality. The Boolean Operator used was "AND" to ensure the presence of the three descriptors in every publication, with the exception of the terms "sound" and "music", for which "OR" was used. Finally, we decided to include book chapters in the review, providing they belonged to books published by publishers catalogued within the SPI (Scholarly Publishers Indicators). There is a large corpus published in this format and it has been considered that it enriches the review.

A total of 252 titles were identified, of which 75 were duplicates, leaving 177 publications. An abstract and keyword query was carried out on these. A selection process was then carried out, excluding a total of 56 articles because they did not meet one of the following criteria: written in English or Spanish (n=13), addressing literacy from the field of communication or education (n=19), being a journal article or a book chapter (n=8), or allowing access to the document (n=16).

During the next phase, the full text of the remaining publications (n=121) was checked, and a final selection process was established. Numerous publications were found that did not present results related to sound or music within the paradigm of multimodal literacy. In many cases, the terms “sound” or “music” appeared only in the definition of the concept of multimodality but were not included in the study itself. Thus, publications were excluded if they did not go in depth into the auditory mode or its relationship with other modes, or if they did not have sufficient data to be able to analyze the use of sound within the study. Publications that did not include multimodal literacy experiences within an educational setting were also excluded during this phase. Literature reviews and multimodal analyses of phenomena that were not framed within educational practices were thus discarded. In this last screening, a total of 73 publications were excluded, resulting in a final sample of 48, 44 articles and four book chapters. We performed an extra search to get some data about the presence of music and sound in the wide corpus of multimodal literacy. The same databases were consulted, using the same Boolean Operators but
omitting the descriptors "sound" OR "music". Knowing that this number would be reduced after a first review, we obtained a total of 1911 results in WOS and 1858 in SCOPUS.

To assist with the analysis, we created a table based on the SPIDER tool (Sample, Phenomenon of Interest, Design, Evaluation, Research type), which is useful for synthesizing data from qualitative studies (Cooke, et al., 2012). The data collected during the reading of the articles are itemized in the Supplementary Table and were classified into the following categories:

- Author and year of publication.
- Type of study in terms of design (action research, case studies, comparative studies, ethnographic studies, etc.).
- Educational stage at which the study was aimed.
- Subject or area of knowledge in which the proposal was framed.
- Role of the student body or study sample.
- Modes of communication involved in the multimodal proposal (aural, verbal, visual, gestural, haptic, etc.).
- Sound elements that materialize the auditory mode: sound effects, music, dialogue, narration, etc.
- Approach to meta-linguistic issues.
- Access to the material.
- Sound domains recognized.

As prior research exists about semiotic implications of sound, we will use a directed approach of content analysis (Hsieh, & Shannon, 2005). Thus, sound domains described by van Leeuwen (1999) will be defined as categories that will be used to analyze the content of the articles and chapters included in the sample. Directed approach in content analysis helps to make the review process more structure and to extend and support the existing theory (Hsieh, & Shannon, 2005).

3. Results

Regarding the year of publication of the papers in the sample, we found the highest frequency in 2017 and 2021, with a total of 7 articles published in each of those years. The fewest publications were found in 2012 (n=1), and 2013 (n=0) (Figure 1). The average number of articles published per year was 4. Many of the articles used similar research designs. We found a significant number that can be classified as "action research" (n=32) and an increasing trend from year 2016 onward. We understand "action research" to mean those studies whose objective is curriculum development or the improvement of educational programmes, in addition to the objectives derived from the research practice itself (Latorre, 2005). Also important, although not as numerous, was the presence of work with an ethnographic character, specially from 2017 to 2021, in which the researcher studied the literacy practices of a community or group (n=12). Some of these ethnographic works were carried out outside the school premises, for example, those starred by a group of young Cuban artists (Butler, et al., 2021), a Filipino community living in the UK (Domingo, 2012), or even the relatives of the researchers who uses autoethnography to assess their multimodal practices (Narey, 2019). Finally, we found one comparative study between different creators of multimodal productions with a series of common criteria (Smith, et al., 2017) and three articles that could be categorized as "case studies" since they involved a detailed observation and in-depth analysis of a specific phenomenon, attempting to analyze all
the factors that influence it (Yin, 2018). It should be noted that all the articles reviewed used data collection instruments and analysis tools typical of a qualitative methodology.

![Figure 1. Number of articles per year of publication and research designs](image)

Regarding the educational stage, we found that 14 of the studies were of secondary school students and 13 of primary school students. These were the two most numerous categories, followed by early years’ education (n=6), higher education (n=6), baccalaureate (n=3) and research conducted beyond a formal educational context (n=3). One of the studies consisted of a training course for graduate education professionals, who were trained in multimodal creations using Scratch (Lemieux, & Manson, 2022). We also found work that involved students’ families (Lozada, et al., 2021), work carried out outside the physical school context (n=3) and one study that brought together students from both primary and secondary education (Hess et al., 2019). In terms of subjects, studies that develop complementary courses in coordination with teachers were the most frequent (n=15). However, the group of subjects was varied: Literature (n=7), English through the Arts (n=5), English as a Second Language (n=5), Music (n=4), and Natural Sciences (n=1). Some of the research occurred in school areas outside the classroom, such as at recess, the playground or on the school bus (n=8), while others occurred in the classroom, but no subject was specified (n=3). This diversity points to a heterogeneous concept of school and education, not always confined to the physical limits of a classroom or the school building. School experience travel to and from other spaces and contexts related to education, merging with the musical and sonic experience of the students. Finally, it is worth mentioning the differences that exist between the different roles adopted by pupils in the studies reviewed. The learner sometimes has the role of creator of the multimodal material, and they are the one who generates the product that usually represents the object of study. This occurs frequently (n=35), while at other times the learner is the receiver of the multimodal message (n=5), analyses a particular resource (n=5) or participates as both creator and analyst (n=3).
We found similarities and differences in the approaches to multimodal literacy (Leander, & Boldt, 2013). Common features included the combination of auditory and verbal modes (n=40). This means that, in these works, auditory material existed alongside elements of verbal language, either spoken or written. It was also common to find auditory material combined with images, logos, gestures, videos, and other visual media (n=35). Sound effects and music were frequently combined with performative or gestural (n=21), spatial and movement (n=10) and haptic elements (n=2). Multimodal proposals that include haptic elements involve students in decoding processes through touching (Dalton, & Musetti, 2018; Stuft, & Gillern, 2021). Furthermore, these modes were presented in up to 15 combinations as can be seen in Figure 2. Notably, we found remediation in 14 studies, although only two of them refer directly to this process, using the term "remediation" itself or some synonym such as "transmodal redesign" (Tomlinson, 2015; Alexander et al., 2016). The auditory mode materialized in very different ways across the sample: sound effects, soundscape elements, narration, onomatopoeias, and rhymes and music, the latter being the most frequent (n=35). We also analyzed the use of metalanguage in the research, i.e., whether the communicative possibilities of the codes used were reflected upon or explored in greater depth. We found that 60.42% of the studies included metalanguage in their approach to the study. However, only 31.35% of the studies allowed access to the material being worked on, always through links or QR codes.

![Figure 2. Frequency of occurrence of communicative modes](image)

We also analyzed the kind of sound domains that appeared explicitly within researchers’ analyses. This does not mean that a song, for example, had no melody, but rather the author has not attributed meaning to melody within the research. We found 10 articles that conveyed meaning by alluding to sonic issues of texture and the interaction of sounds and voices. One example is a study in which students created a podcast exploring the relationships that exist between the different sounds they found at school (Doerr-Stevens, & Buckley-Marudas, 2019) or a study that analyzed the sounds in the playground and in some primary school children’s games (Countryman, & Gabriel, 2014). We identified 13 studies in which reference was made to melodic
or pitch elements. Burn (2016), for example, analyzed the impact of descending scales or an ostinato in the production of animation movies by his students. In terms of sound, 25 of the works included elements related to directionality or intensity in some way. Some researchers used virtual reality (Hutchison, 2018; Yeh, & Tseng, 2020), digital sound effects (Stufft, & Gillern, 2021), and sound walks (Wargo, 2018b). The research that investigated temporal or rhythmic sound elements (n=24) were often related to the prosody or rhythm of the text (Höglund, 2022), while in other cases they were used in musical productions such as rap (Morgade, et al., 2016). The timbre domain, on the other hand, was found in works that assessed the sound of different voices or instruments, which appeared in a total of 21 research items. It is interesting to note the relationships established between the different domains: frequency of groupings, possible incompatibilities, hierarchies, etc. Table 1 shows the frequency of appearance of the sound domains in each of the articles.

**Table 1. Sound Domain Frequency and Correlation**

<table>
<thead>
<tr>
<th>Sound Domains (van Leeuwen, 1999)</th>
<th>Voices interaction</th>
<th>Melody</th>
<th>Modality</th>
<th>Perspective</th>
<th>Time</th>
<th>Timbre</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alexander, et al., 2016</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baize, 2019</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barton, et al., 2022</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blom, 2017</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burn, 2016</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Butler, et al., 2021</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen, 2021</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Countryman, &amp; Gabriel, 2014</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Dallacqua, et al., 2015</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dalton, &amp; Musetti, 2018</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dalton, &amp; Smith, 2015</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doerr-Stevens, &amp; Buckley-Marudas, 2019</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domingo, 2012</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flint, et al., 2021</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>García, et al., 2021</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hackett, et al., 2017</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harrop-Allin, 2017</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hess, et al., 2019</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Höglund, 2022</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huang, 2017</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hutchison, 2018</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim, &amp; Li, 2021</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lemieux, &amp; Mason, 2022</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lewis, 2020</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lim, &amp; Tan, 2018</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lozada, et al., 2021</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Finally, the modality domain refers to the figurative capacity of sound. To define this category, we evaluated whether the "iconicity" of the sound event was analyzed in the study, which might make it possible to make a classification between a greater or lesser modality, that is, a greater level of closeness between the sound and its real manifestation, between signifier and signified. Modality is a quality that encompasses, in a way, all the other domains. All elements such as time, perspective, pitch, etc. influence the ideas we construct about sound (van Leeuwen, 1999). In the work of Alexander et al. (2016), for example, a student plays with this modality intention, modifying sound effect recordings that a priori seem unconnected to the visual material these sounds are presented with. It should be noted that those articles that incorporated musical elements specifically addressed melodic, rhythmic, or texture-related issues (Tomlinson, 2015; Williams, 2020). However, in other cases (Dalton, & Musetti, 2018), the song is understood as a whole and the impact of music as a compendium of elements is reflected upon. On other occasions, the articles focused on the auditory analysis of the lyrics of a song, an element that would not fit within the domains proposed by van Leeuwen (Baize, 2019). We found research where, albeit considering sound and music as modes, were primarily focused on more traditional understandings of multimodality. For instance, Kim and Li (2021) mentioned music as one of the multiple resources that students effectively utilized in their compositions. However, no sound domain or music trait were mentioned in the analysis of these products, and their analysis dwells mostly on the visual or writing modes.
5. Conclusions

Music (and sound) is an abstract language with a great capacity to communicate, transmit and represent (DeNora, 2009). The digital transformation of society has allowed new ways in which this language interacts with other disciplines, and the multimodal conception of literacy opens up a multitude of possibilities for its application in education (Salmon, 2010). However, relatively few researchers choose to focus on acoustic issues when developing multimodal literacy research designs. Going back to the results obtained in the identification phase of the review in this paper (110 in WOS and 143 in Scopus) and comparing them with the results obtained with the descriptors multimodal* AND literac* (1911 in WOS and 1858 in Scopus), we can conclude that there is a very high percentage of studies that exclude or do not consider sound or music within their multimodal literacy approach. These findings align with the need identified by Elwick et al. (2020, p. 179) to “design and advocate for research that meaningfully engages music and sound practice and discourse and, in doing so, contributes to a new understanding of children’s experiences of music and sound practices”.

According to our first objective, this review has allowed us to consider the trends in studies that deal with sound within multimodal literacy, and also to highlight the gaps that still exist in this line of research. We have detected a tendency to carry out ethnographic studies at early educational stages (early childhood and primary education) and to develop action research in which students create or analyze multimodal elements at secondary or higher education level. It would be interesting to study in greater depth the way in which the role of creator is developed in early years’ education and, in turn, to use ethnographic tools to analyze adolescent literacy practices. The need for research in this respect makes particular sense if we take into account the number of multimodal resources that can be found in social networks, video games and other platforms used regularly by young people. On the other hand, the lack of quantitative studies in the corpus emphasizes a typical trend in music education research in which the sample is usually small-scale. The use of qualitative methodologies also shows the multiplicity of multimodal practice, hard to be depicted in academic research through quantitative methods. Artistic experience is considered relational and experiential, and research should attend to all agents that may influence how children create their aural worlds (Rowsell, 2013; Wargo, 2018b).

Secondly, regarding the semiotic implications of the auditory mode in multimodal research, we can say that sound has a high versatility, as a multitude of sound elements are identified in the different studies; these include sound effects, vocalizations, chants, narrations, musical elements, songs and so on. In line with Nash (2020), we conclude that sound can have a high level of narrative impact. However, we have found several studies that limit the auditory mode to the use of "background music", without making explicit the relationship between the music used and the communicative intent of the message or meaning (Baize, 2019; Blom, 2017; Dalton, & Musetti, 2018). When limiting the use of music and sound to the background, as a mere accompaniment, their capacity of communicating and thrilling is neglected, sometimes motivating the receiver to “ignore” the aural inputs (DeNora, 2009). On the other hand, articles such as Burn’s (2017) make a complex and complete use of the auditory mode, giving weight within the message to different sound elements. We have also seen that most studies have a representational approach, based on fixed relationships between signifier and signified (MacLure, 2016). We advocate broadening the field of study around the sensory by focusing on the listening or creating experience itself, and on the relationships established between sound and participants.
Powell, & Somerville, 2020). Recent interpretations of sound as a holistic experience shows a much deeper multisensorial complexity, but sound has not explored yet its semiotic potential beyond van Leeuwen’s work in 1999. The moment we move away from the linguistic to focus on embodiment, the possibility opens up to understand the experience of sound in a more integrated and complex way.

Analysis of the data allows us to venture that new technologies will encourage the emergence of new sound elements, as has been happening in recent years with, for example, virtual reality (Hutchison, 2018; Yeh, & Tseng, 2020). This may provide an opportunity for researchers and educators to design innovative multimodal approaches that explore acoustic possibilities to their fullest. Our review aims to serve as a guide for teachers interested in multimodal literacy to develop innovative educational practices in this field. It has been found that multimodality often leads to a lexical and visual approach, while some authors defend the possibilities of sound in multimodality (Wargo, 2018a). This review challenges the limitations to which the concept of multimodality can lead to, as it was conceived in a sort of “aural marginalisation”. In addition, it is intended to lay the foundations for a solid line of research that will facilitate a further exploration of multimodal literacy from the perspective of sound studies, with the intention of recognising and expanding the communicative and educational possibilities of sound. Despite results show a hopeful future for sound in multimodality, much work should be done to equate sound and music with visual and verbal in multimodality.

Funding and Acknowledgement

This research is part of the R&D Project PID2019-104557GB-I00, funded by the MCIN/AEI/10.13039/501100011033/, and the R&D Project P20-00487, funded by the European Regional Development Fund (ERDF) and Junta de Andalucía. Special recognition to Universidad Loyola Andalucía, in gratitude for the Research Assistant grant that made possible this research.

References

Those titles preceded by an asterisk (*) are part of the review corpus:


*Park, J.H. (2021). "Dear Future Me: Connecting College L2 Writers' Literacy Paths to an Envisioned Future Self Through a Multimodal Project. In D. Shin, T. Cimasko, & Y. Yi (Eds.), Multimodal Composing in K-16 ESL and EFL Education (pp.73-86). Springer Singapore. [https://doi.org/10.1007/978-981-16-0530-7_5](https://doi.org/10.1007/978-981-16-0530-7_5)


*Yeh, H.-C., & Tseng, S.-S. (2020). Enhancing multimodal literacy using augmented reality. *Language Learning, 24*(1) 27-37. [https://doi.org/10125/44706](https://doi.org/10125/44706)