

PROSODY IN TOURIST ORAL COMMUNICATION Pausing in city audio guides

MARIA ELISA FINA
UNIVERSITÀ DEL SALENTO

Abstract – Over the last decades, tourist communication has been widely investigated in various languages, involving a variety of printed and Web-based genres and a variety of perspectives. However, as far as oral tourist genres are concerned, there seems to be a gap in academic research, with the published literature including a study by Gavioli (2015) on dialogue interpreting in guided tours in English, a study by Francesconi (2014) on prosody in a BBC travel radio programme, a study by Ravazzolo (2014) on guided tours of botanic parks in French, and a study by Nardi (2012) on museum audio guides in German. Even lesser attention has been paid to the prosodic features of speech in oral tourist genres, and how these are employed to effectively deliver the contents of the tour. This study attempts to contribute to developing this specific area of tourist studies by investigating prosody in city audio guides. More specifically, the study focuses on the use of pauses in city audio guides in Italian and in English. At a theoretical level, the city audio guide is intended here as a form of informative script-based public speaking, whose main function is to provide historical and artistic information on the sites visited by users. On the basis of such theoretical premises, the study unfolds around the assumption that prosody – and pausing in this specific case – plays a key role in effective content delivery and uptake. The study was carried out on 42 audio fragments selected from the scripts of 21 audio guides produced in Italy, the UK and the USA. In each audio fragment, silent pauses were annotated by means of *PRAAT*, a software for acoustic analysis of speech. The use of pauses between utterances was then investigated in each group of fragments, quantitatively in terms of pause duration variation, and qualitatively in relation to the textual features of the selected portions of scripts. The results show a more frequent use of long pauses and patterned pausing in the British audio fragments with respect to their Italian and American counterparts. Finally, the results are framed within a discussion on best practices in public speaking.

Keywords: city audio guides; prosody; pausing; public speaking; acoustic analysis.

1. The ‘great gap’ in Italian/English tourism studies

Over the last decades, tourism communication has been widely explored in academic research, and a wide range of tourist genres have been thoroughly analysed by combining a variety of cross-disciplinary perspectives and methods. In the Italian context, great attention has been paid to tourism discourse in Italian and in English, with research moving progressively from traditional printed genres such as guidebooks, brochures and magazines (Cappelli 2012; Denti 2012; Francesconi 2011a; Katan 2012; Maci 2013), to Web-based genres such as websites (Cappelli 2008a; Maci 2013, Manca 2013, 2016; Nigro 2012; Plastina 2012), including user-generated content such as travel blogs (Cappelli 2008b; Denti 2015), travel reviews (Cappelli 2013; Fina 2011), travel articles and travel reports (D’Egidio 2013, 2014), to audiovisual products such as promotional videos (Francesconi 2011b, 2014, 2015),¹ with the perspectives of investigation ranging from contrastive discourse analysis and intercultural communication, to issues related to

¹ Only the most recent studies have been quoted.

translation, remediation and translation quality, to meaning making by means of multimodality.²

While so far academic research has been focusing on written communication and the printed or moving image, very little research has been carried out on oral tourist genres in Italian and in English, such as guided tours, audio guides, travel television programmes, and similar. Gavioli (2015) carried out a study on dialogue interpreting in guided tours from Italian into English, and she found out that interpreters tend to expand the guides' presentation, adding information that the guides seem to take for granted (e.g. familiarity with specific terms, units of measures), clarifying concepts or asking questions that make particular details relevant. However, she did not investigate the prosodic component of the interpreter's speech.

The prosodic dimension was investigated by Francesconi (2014), instead, who analysed prosodic features of speech in an episode of a BBC Radio 4 travel series. She noticed a systematic segmentation in intonation units even in very simple sentences, fast speech, with the effect of conferring dynamism and energy to speech, the use of pitch accent to mark linguistic items corresponding to particular artistic and architectural elements, and variation of loudness from soft to loud. To the best of my knowledge, no further literature is available on the prosodic component of speech in tourism communication in Italian and in English. Hence, the aim of this study is to contribute to developing this unexplored area of tourism studies by investigating prosody in city audio guides.

Over the last decades, city audio guides have been produced and made available on the Internet or as apps to satisfy the preferences of independent travellers wishing to sightsee on their own without the constraints involved in human-guided tours. Despite the huge amount of audio guides produced in various languages worldwide, this tourist genre has become object of academic research only very recently. Great attention has been paid mainly to museum audio guides, with particular reference to systems for enhancing usability (see, for example, Alfaro *et al.* 2005; Gebbensleben *et al.* 2006; Petrelli and Not 2005) and accessibility (Neves 2012, 2013). Audio guides have also been investigated from a translation perspective, with a study by Tempel and ten Thije (2010) on the appreciation of cultural and linguistic adjustments in English and German translations of Dutch museum audio guides, but, to the best of my knowledge, no investigation on the prosodic features of audio guides has been carried out so far.

The study presented in this paper focuses on the use of pauses in a corpus of city audio guides, as part of a wider cross-language, multimodal investigation of city audio guides in Italian and in English (Fina 2016a) in all their verbal and extra-verbal aspects.³ From the methodological point of view, the study draws on a number of previous studies on different spoken genres, which will be illustrated in the following section.

2. Relevant studies

The research carried out in this paper draws on a number of studies that, though treating other spoken genres, are relevant to the purpose of my investigation and can be considered

² For a full account, see Peverati (2012) and Francesconi (2014, pp. 166-172).

³ See also Fina 2016b.

valuable starting points for developing an adequate methodological approach to the analysis of pauses in the audio guides.

One of the main reference studies is that by De Meo *et al.* (2015), who investigated the impact of supra-segmental features of voice on multimedia distance learning. The voices selected for the experiment included voices of Italian native speakers (both neutral and regionally marked), voices of non-native speakers (English and Chinese), and synthesised voices with the prosodic traits of the natural voices. A scientific text was read by each voice to 165 students in the *Campania* region, who were then given a comprehension test based on word/picture matching and sequence re-ordering. A comparison of the results of these tests with the prosodic traits analysed in each voice (articulation and speech rates,⁴ pitch range and speech composition) showed that the Chinese voice had the most negative effects on the comprehension tests. In the acoustic analysis, this voice was characterised by a slower articulation rate, more frequent silent pauses and narrow pitch range, which are obviously due to a lower mastery of the Italian language with respect to the other speakers. In contrast, the native speakers' voices (including those regionally marked) proved to have positive effects on the understanding of the text, and are characterised by faster articulation rates, fewer pauses and broader pitch ranges.

Relevant to this study are also a number of studies by Strangert (1991, 1993, 2005), in which she investigated professional news reading, political spontaneous talk and similar genres in Swedish. Particularly interesting is her investigation of the distribution of pauses, and of their acoustic correlates, in read-aloud news texts (1991), since she combined prosodic, syntactic and textual perspectives by investigating pauses in relation to text structure. The author distinguished between 'strong pause positions' and 'weak pause positions', and identified length, information load and semantic load as the key factors governing the distribution of pauses. In a similar study (1993), she also noticed that in professional news reading demands for effectiveness and intelligibility are satisfied by means of specific strategies, in which pauses occur at syntactic boundaries, but also seem to be semantically governed, since they tend to occur before the most salient words. In another study (2005), Strangert also looked closely at pitch dynamics, and noticed that the above-mentioned pausing patterns tend to co-occur with extensive use of pitch range and high peaks to put extra weight on important segments of speech.

Although the audio guide genre is very different from the spoken genres analysed in the above-mentioned studies, most of these have provided important methodological insights into the investigation of prosody, which, in part, have been applied to my research. Before illustrating the data and methodology, it is necessary to theoretically discuss the stylistic nature of the audio guide.

3. The hybrid stylistic nature of the audio guide genre

The investigation of prosody in city audio guides is motivated in this study by theoretical observations about the stylistic nature of the audio guide as a text type, which will be illustrated by referring to the stylistic discourse categories developed by Crystal and Davy (1969). The two scholars distinguish between two stylistic categories of discourse: *medium*, in which features distinctive of speech vs. writing are grouped, and *participation*,

⁴ The terms 'articulation rate' and 'speech rate' both indicate the pace at which speech segments are produced, with the difference that speech rate includes pause intervals, while articulation rate does not.

in which features distinctive of monologue vs. dialogue are grouped. In particular, they investigated the category of ‘medium’ by looking at the distinctive features of speech and writing in reference to the classification of texts in the *Survey of English Usage* (SEU)⁵ (Quirk 1959), which is based on the fundamental distinction between speech and writing. The two scholars focused their attention on the “undesirable asymmetry” (Crystal 1994, p. 36) arising from the fact that not always does language stay in one category (either written or spoken), but there are cases in which switch occurs as, for example, in dictation, where speech is produced to be written down.

This is also the case of the audio guide text which, as can be seen in Figure 1, at a first stage comes as a written script containing the description of the landmarks of the city. At a second stage, this script will be read aloud by a professional speaker, and will be recorded for subsequent publication:

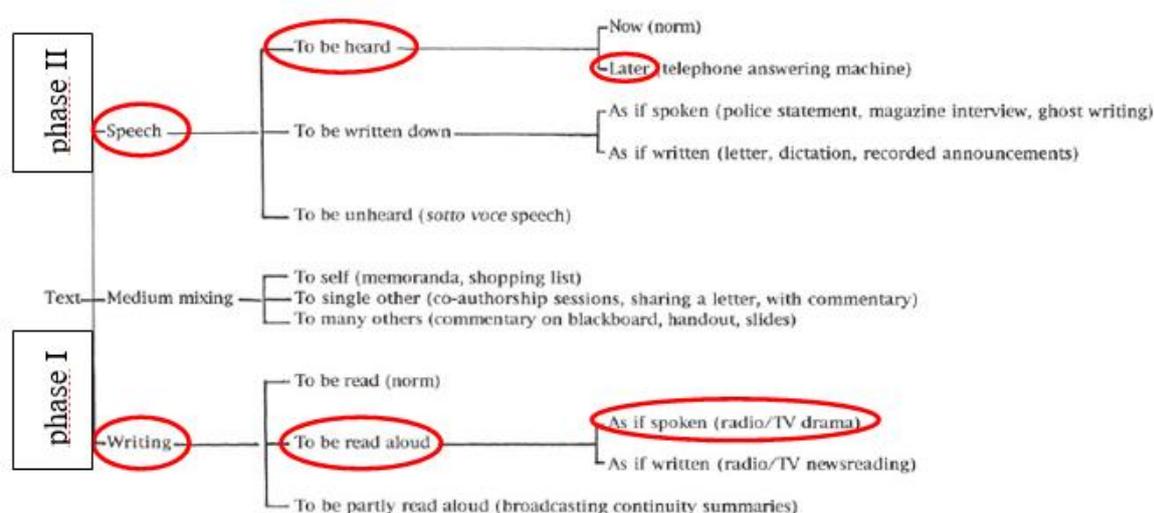


Figure 1
Classification of the audio guide as a complex medium text.

Thus, the audio guide text is characterised by what Crystal and Davy term ‘complex medium’ to classify “certain features of a variety which would fall as a general rule within one kind of discourse by reference to features which would normally be expected to occur only in another” (1969, p. 71). This double stylistic nature places the audio guide halfway between written and oral, and classifies it as *trasmesso* (Sabatini 1997), i.e. ‘broadcast’.

On the basis of these theoretical observations, the audio guide can be considered a particular type of public speaking genre, i.e. an informative, script-based oral presentation. As such, the hybrid stylistic nature of the audio guide certainly has implications in terms of content delivery and uptake. A wider investigation of contents in the audio guides (Fina 2016a) showed that the scripts are informative texts dense with information about the historical and artistic features of the sites. In other words, audio guides consist of texts that are delivered orally, but that bear the complexity of written texts. This makes the crucial issue of oral delivery particularly significant, since the audio guide will be fully accessible and enjoyable for users only if oral delivery takes place by adopting effective prosodic

⁵ <http://www.ucl.ac.uk/english-usage/about/history.htm> (last accessed February 2016).

strategies that allow maximum content uptake with minimum cognitive effort. In the following sections, we will investigate pausing as one of these possible strategies.

4. The study

4.1 Data

The data used for prosodic analysis consist of 42 audio fragments selected from 21 original⁶ audio guides produced in Italy, the UK and the USA,⁷ as reported in Table 1:⁸

Italian audio guides	British audio guides	American audio guides
Bologna	Bristol	Chicago
Genova	Cambridge	Denver
Matera	Edinburgh New Town	Greenwich Village NYC
Milano	Edinburgh Royal Mile	Little Rock Arkansas
Padova	London City	London by Rick Steves
Perugia	Northampton	Washington National Mall
Torino	Portsmouth	Philadelphia

Table 1
Audio guides selected for the analysis.

As can be seen, each group includes audio guides covering cities belonging to the corresponding country, so as to have audio guides of Italian cities in Italian, audio guides of British cities in British English, and audio guides of American cities in American English. The only exception is the audio guide of London produced by the American travel expert Rick Steves, who is considered a leading authority in the production of audio guides. The selected audio guide is of a British city because Steves produces audio guides of European destinations only,⁹ and this made it impossible to choose a tour of an American destination. However, the fact of including an audio guide of a British city produced by an American does not affect the research because for prosodic analysis the voices need to be classified according to the speakers' nationalities. For convenience, we will refer to 'Italian', 'British' and 'American Corpus' when addressing the audio fragments of the respective groups.

For each audio guide, two music-free audio fragments, lasting around one minute each and produced by male voices, were selected from two different sections of the scripts. In all the audio fragments, the contents are dense with historical information and/or artistic description. Whenever possible, I tried to balance narration and description either within the same audio fragment or between the two audio fragments of the same audio guide, so as to avoid excessive discrepancies in terms of content types. Furthermore, whenever possible, descriptive sections focusing on the same type of element (e.g. the façade of a building) were selected.

⁶ By 'original' we mean that the original languages of the selected audio guides are Italian, British English and American English respectively.

⁷ This feature was retrieved by looking at the address of the producers in each of the website.

⁸ See *References* for a list of the websites.

⁹ See <https://www.ricksteves.com/watch-read-listen/audio/audio-tours> (last accessed February 2016).

4.2 Method

Speech analysis in the selected audio fragments was carried out by means of *PRAAT*,¹⁰ a free scientific software for acoustic analysis of speech. First of all, the content of each audio fragment was divided into utterances, and each utterance was then divided into numbered phonetic syllables. Phonetic – rather than orthographical – syllabification implies a more accurate annotation, since only the syllables actually produced by the speakers were annotated. Silent pauses were annotated both between utterances and within the same utterance (from now on utterance-internal pauses). Expectedly, since the type of speech we are dealing with is script-based and hence highly controlled, no hesitation-filled pauses could be detected. For each audio fragment a *PRAAT* textgrid was created, and by means of an algorithm the duration of silent pauses was retrieved, and mean duration of pauses, along with standard deviation values¹¹ (from now on ‘st.dev.’) were extracted.

Following studies by Strangert (1991, 1993, 2005), the analysis was carried out by referring to the textual interrelations that make the selected portions of text cohesive and coherent. More specifically, the use of pauses between utterances will be investigated at the discourse level, by referring to the various topics and sub-topics into which the text may be divided. In a wider investigation of the audio guides (Fina 2016a), both pauses between utterances and utterance-internal pauses were investigated, but due to space constraints, in this paper we will focus on pauses between utterances only.

4.3 Analysis

Pausing in what Strangert (1991, p. 239) defines “strong pause positions”, i.e. at utterance boundaries, is extremely important in audio guides because, if done effectively, it allows the listener to receive information in appropriate doses, depending on the type and amount of contents of the single utterances. What we are interested in is to investigate the distribution and duration of pauses between utterances in relation to text structure and content, in order to identify specific patterns and/or differences between the Italian, British and American audio fragments.

The mean duration values and st.dev. values were examined in the audio fragments and compared across the three corpora. The data reflect a high degree of diversification not only between the three corpora but also within the same corpus. The mean duration values were grouped into categories defining the ranges of mean pause duration for each corpus, as reported in Table 2:

	ITA	UK	USA
Ranges of mean duration	0.5s ¹² – 1.8s	0.8s – 1.8s	0.4s – 1.2s
St.dev.	0.6s	0.5s	0.3s

Table 2
Pauses between utterances: mean duration ranges and st.dev.

¹⁰ The software has been designed and developed by Paul Boersma and David Weenik at the University of Amsterdam, and is available at <http://www.fon.hum.uva.nl/praat/> (last accessed February 2016).

¹¹ In statistics, standard deviation is a measure indicating “the amount by which measurements in a set vary from the average for the set”. (*Oxford Advanced Learner’s Dictionary*).

¹² 0.1s corresponds to 100 milliseconds.

In line with the mean proportion of silence identified in each corpus, while in the British audio fragments the mean duration is generally high, ranging from 800ms to almost 2s, in the Italian and the American audio fragments we may also find shorter pauses and, considering that these are mean values, in single fragments we may also find pauses below 400ms. The values of pauses in the Italian and the British audio fragments are also associated with higher st.dev. values, which generally indicate high variation in pause duration. What we are interested in is to find out whether high st.dev. values are due to variation in pause duration between the various audio fragments or within audio fragments and, in the latter case, whether variation is patterned in relation to textual features.

The position at which long pauses occur were examined in the three groups of audio fragments, and it was observed that pauses long 1s or more tend to occur before a new item is introduced, or to signal the beginning of a sub-topic of the main topic. In the following examples from the British Corpus, pauses long more than 1s occur before a new item is introduced:

- (1) The overall effect has been described as Moorish in appearance. (1.8) Behind it, a twelfth-century inner porch from the earlier church still stands [...] (Bristol, audio 1, min. 00:29)
- (2) The statue is of Adam Smith, the famous eighteenth-century Scottish economist, and author of *The Wealth of Nations*. (1.5) Across the road from the Mercat Cross, there used to be the house of Sir Simon Preston, the Provost of Edinburgh in 1567. (Edinburgh Royal Mile, audio 2, min. 00:39)
- (3) This [the “Whispering Gallery”] is also the best position to view the famous Thornhill frescoes. (2.0) After another 120 steps, if you are not out of breath by now, you will find the “Stone Gallery”. (London East End, audio 2, min. 00:19)

In all three examples, the visitor’s attention is focused on a different item. While in (1) and (3) the new items introduced (i.e. “the twelfth-century inner porch” and “the ‘Stone Gallery’” respectively) are part of the same place, in (2) a new building is indicated. In all three cases a change in direction or in item location is involved, as marked by the clauses “Behind it” in (1), “Across the road” in (2) and “After another 120 steps” in (3), which purposely occur at the beginning of the utterance.

Another particular example is the following:

- (4) It [the Guildhall] was extended in 1892 by Matthew Holding, and a more recent sympathetic addition was built to the east in 1992. (1.8) Cross the road carefully and look at the ornate carvings above the first-floor windows. (Northampton, audio 1, min. 00:18)

In the section preceding the pause, the history of the development of the Guildhall is narrated from across the road. After having listened to the history of the site, the visitor is then invited to cross the road and look closely at the site to identify the “ornate carvings”. Thus, in this case the pause not only marks the end of preliminary information about the site, but also involves a change in the physical perspective from which the item is observed.

A long pause can also be used to signal the end of a subtopic and the beginning of another one, both related to the same site, as in the following example:

- (5) Woodelarke, a deeply religious man, ordered that students should study only philosophy, or theology. (2.5) There are two theories regarding the name of the college. (Cambridge, audio 1, min. 00:19)

In terms of information structure, the two utterances divided by the pause constitute two unrelated subtopics (the founder and the name of the College respectively) of the same

topic (St Catharine's College). Thus, the cognitive function of the pause is evident, as the long pause signals that a new topic is going to begin, and at the same time allows the visitor to fully absorb the previous information before going on with the narration.

A case of particularly functional pausing is the following:

- (6) This [the trial of Thomas à Becket] is above the last window to the left of the entrance as you face the building. (1.8) Becket was summoned to Northampton Castle to defend his opposition to new legislation, intended to allow the King's court to try members of the clergy. (Northampton, audio 1, min. 00:45)

The long pause here gives time to the visitor to identify the item, i.e. a carving representing the trial of Becket, before starting to explain the historical meaning of the decorative piece. A similar example can be found in *Bristol*, with a long pause occurring after item presentation:

- (7) Standing outside this imposing building, one can appreciate why Queen Elizabeth I called it "the goodliest, fairest and most famous parish church in England", on her visit to Bristol. (2.5) You enter by the north porch, which has an interesting seven-pointed archway and was built in the early fourteenth century. (Bristol, audio 1)

The long pause after the first utterance is meant to give time to the visitor to take in the effect of the "imposing building" and to recognise in it its qualities, boasted through the words of Queen Elizabeth I.

Finally, utterances containing information on the same item or topic may also be aurally distanced by pauses of exactly the same duration, as in the following examples:

- (8) The crypt or lower church dates from much earlier, and is an elegant building of the early fifteenth century. (1.2) It contains the tomb of John Whitson, merchant and founder of the Red Maids' school in Bristol, still the oldest foundation for girls in the country. (1.2) Every year, the school holds a memorial service at the tomb in the presence of the Lord Mayor. (Bristol, audio 2, min. 00:41)
- (9) Saint Catharine's College was originally founded in fourteen seventy-three by Robert Wodelarke, who was provost of King's College. (1.4) At that time, it is alleged, there were suspicions that Wodelarke was illicitly taking funds from King's College to fund the building of Saint Catharine's. (1.4) Wodelarke, a deeply religious man, ordered that students should study only philosophy or theology. (Cambridge, audio 1)
- (10) The building on the corner, off the north side of the square, which is now used by the IBM company, was the birthplace of Henry Peter Brougham, the first Baron Brougham and Vaux. (1.0) He was an orator, lawyer and politician, also a legal reformer and one of the prime movers of the institution of the University of London. (1.0) He was a noted member of William Wilberforce's Clapham Sect, the anti-slavery movement whose 200th anniversary we celebrated in 2007. (Edinburgh New Town, audio 1)

In these cases, the steady duration of pauses gives rhythm to the narration, but at the same time allows the listener time to process the information. The examples presented so far show that in most British audio fragments the variation in pause duration, which is reflected in high st.dev. values, tends to be distributed according to content-related criteria.

However, high st.dev. values do not necessarily imply a patterned variation in pause duration. In most Italian audio fragments, indeed, pauses are generally below 1s, and no great variation occurs. This means that the high st.dev. value detected in the Italian Corpus – which is even slightly higher than the value in the British Corpus – is due to high variation of pause duration *between* the selected audio guides rather than within the same audio fragment. This is confirmed by mean duration values retrieved for each audio

fragment, which are reported in Table 3:

Audio guide	Fragment n.	Mean duration (s)
Bologna	audio 1	0.4
	audio 2	0.6
Genova	audio 1	0.6
	audio 2	0.9
Matera	audio 1	1.7
	audio 2	1.9
Milano	audio 1	1.1
	audio 2	1.1
Padova	audio 1	0.9
	audio 2	0.8
Perugia	audio 1	0.7
	audio 2	0.9
Torino	audio 1	1.4
	audio 2	1.8

Table 3
Mean duration values of pauses in the Italian Corpus.

As can be seen, the audio fragments display very different mean duration values, which could range from less than 0.5s to almost 2s. The use of long pauses with content-related patterns can be found in the audio fragments of *Matera* and *Milano*:

- (11) *Diamo le spalle al Sasso Barisano e ammiriamo gli ornamenti di questa “bella Signora”. (3.0) Gli archetti del coronamento sorretti da dodici pilastri: gli Apostoli, che seguono Cristo con la sua Croce posta in alto. (1.1) Sotto, altri quattro pilastri laterali: gli Evangelisti, testimoni della storia della Salvezza. (Matera, audio 1, min. 00:03)*
[Let’s leave the *Sasso Barisano* behind to admire the decoration of this “beautiful Lady”. The bows of the pediment sustained by twelve pillars: these are the Apostles, who follow Christ, with his Cross placed above. (1.1) Below, other four lateral pilasters: these are the Evangelists, who witnessed the history of Deliverance.]
- (12) *Furono realizzati nel Seicento su disegno del Cerano, insieme alle relative fasce di bassorilievi e alle finestre del primo piano. (1.5) Delle cinque porte in bronzo, la più recente è la prima da destra, opera del 1965 di Luciano Minguzzi, il quale vi illustrò episodi della storia del Duomo. (1.3) La più antica è quella centrale, decorata alla fine del milleottocento con i rilievi di Ludovico Pogliàghi, raffiguranti storie della vita di Maria. (Milano, audio 1, min. 00:06)*
[They were built in the 17th century on Cerano’s design, along with the corresponding series of low-reliefs and the first-floor windows. (1.5) Of the five bronze gates, the most recent is the first one on the right, built in 1965 by Luciano Minguzzi, who decorated it with illustrations representing the history of the Cathedral.]

In both cases, long pauses are used after descriptive sections to give time to the visitor to observe the items indicated in the utterances preceding the pauses. It can also be noticed that shorter pauses are used to shift the attention on different items of the same ‘whole’.

As in the British audio segments, long pauses may also occur when a shift in focus involves a shift in direction or location as well:

- (13) *Il titolare della chiesa, San Nicola, si distingue benissimo sul fondo dell'abside di sinistra, accompagnato da Santa Barbara e San Pantaleone. (2.0) Nel presbiterio destro, una Cristo sulla croce indica alla Madre il "discepolo che Egli amava", quel San Giovanni che ne annuncerà anche l'evangelico ritorno alla fine dei tempi. (Matera, audio 2, min. 00:19)*
 [The church's title holder, St. Nicholas, is well visible in the background on the left apse, together with St Barbara and St Pantaleone. (2.0) In the right chancel, a Christ on the Cross shows to the Holy Mother the "disciple He loved", that St John who would announce His return until the end of time.]
- (14) *La cripta di fronte fu invece disegnata da Pellegrino Tibaldi, ed è una splendida cappella circolare, ricca di stucchi e marmi colorati, in cui sono custodite, sotto l'altare al centro, le reliquie dei martiri milanesi. (1.4) Di nuovo in superficie, si prosegue attraversando la chiesa per raggiungere il transetto sinistro, suddiviso come quello di fronte in tre navate. (1.1) In fondo alla destra c'è una piccola porta, ora chiusa al pubblico, che immetteva alla Scala dei Principi, un tempo riservata solo alle più alte personalità. (Milano, audio 2, min. 00:13)*
 [The opposite crypt was designed by Pellegrino Tibaldi, instead, and is a beautiful circular chapel, richly decorated with stuccoes and coloured marble. Below the altar in the middle, it contains Milanese martyrs' remains. (1.4) Back on the surface, we proceed across the church to the left transept, which is subdivided into three aisles, as the one on the opposite side. (1.1) All the way down on the right-hand side there is a small door – now closed to the public – that led to the *Scala dei Principi*, which could be accessed only by the high personalities.]
- (15) *Però, oltre agli elementi neoclassici ci sono anche i grifoni neogotici, che adornano le terrazze; e poi, i leoni egizi a terra. (1.0) Il piano superiore, che è visitabile, ha una sala ercolana, una sala egizia, una sala cinese, con arredi e decori in stile. (Padova, audio 2, min. 00:14)*
 [However, in addition to neoclassical items there are also the neogothic griffons decorating the terraces, and then, the Egyptian lions on the floor. (1.0) The upper floor – which is accessible – has a Herculaneum room, an Egyptian room, a Chinese room, with furniture and decoration in the same style.]

In these examples, pauses around 1s long are used when the visitor is required to change direction, i.e. from left to right in (13), and from the crypt basement to the surface and then to the right in (14). In excerpt (15), instead, the pause rather marks the end of the description of the exterior, since the visitor is not actually required to physically move to the first floor.

As far as the other Italian audio fragments, no significant pausing patterns were noticed. In those of *Bologna*, *Perugia* and *Genova* the pauses are generally below 1s, whereas in those of *Torino* long pauses occur at each utterance boundary, without any specific pattern.

In the American audio fragments, the low st.dev. value reflects the generally stable duration of pauses between utterances. As in the Italian Corpus, in most audio fragments pauses are below 1s, and no specific criteria could be detected in the distribution of pauses long 1s or more. However, isolated cases of content-related pausing were found, as in the following example:

- (16) Looking at the Michigan Avenue façade, you see the three arched entranceways to what used to be the Auditorium's 400-room hotel. (0.4) It's now Roosevelt University. (1.2) Throughout his career, Sullivan was trying to create a uniquely American architectural style. (0.4) He rejected the popular historical style such as Neoclassicism, that borrowed elements from ancient Greece and Rome. (Chicago, audio 1, min. 00:14)

Although the main topic is the same, i.e. the Auditorium Building in Chicago, the long pause is intended to mark the boundary between two different aspects of the same site. In the first part the speaker illustrates the physical aspects of the building (i.e. the "three arched entranceways"), while in the part following the 1.2s-long pause the description turns to the architect and his artistic ideals. In contrast, short pauses of the same duration are used between utterances strictly interrelated in terms of topic.

Content-related pausing can be also found in the following excerpt:

- (17) The main floor, now the bar and restaurant, served as the original JS Brown Mercantile showroom in 1899, while the second floor now houses an excellent poolroom. (1.1) In the basement is a comedy club and is also where the large brew kettles are housed. (0.6) Wynkoop Brewing offers a great selection of food as well as several beers on tap that are brewed on-site. (0.7) If you wander inside, note the original oak trim and maple floors. (0.5) The pressed tin ceiling on both levels is not only decorative, but also was used to deflect sound down from the lofts that are on the upper stories of the brewery. (Denver, audio 2, min. 00:19)

Here, the change in focus also implies a change in direction, with the long pause marking a change from the “main floor” into the “basement” of the site. Afterwards, instead, shorter pauses lasting averagely 600ms are used to distance the indication or description of decorative items. It might be argued that, if the visitor is inside, 500-700ms may not be enough to allow the visitor to observe the items. However, an investigation of content types on a wider corpus of audio guides (Fina 2016a) revealed that in most British and American audio guides the interior of buildings tends not to be extensively discussed, and only the most important features are mentioned by the speaker in case the visitor wishes to go inside. This is exactly what happens in (17), as indicated by the dubitative clause “if you wander inside”, meaning that the visitor is still outside and the speaker is simply listing interesting items to look at once the visitor is inside.

A content-related pattern can also be found in *Greenwich Village NYC*:

- (18) The impressive structure in front of you is the Washington Square Arch. (1.3) This impressive Roman style arch designed by Stanford White is one of the first expressions of the City Beautiful Movement in New York. (Greenwich Village NYC, audio 1)
- (19) The current version was completed in 1891 in white marble, costing 133,550 dollars. (2.3) The white marble structure features two statues of Washington on the north face, Washington as Commander-in-Chief Accompanied by Fame and Valor [...] (Greenwich Village NYC, audio 1, min. 00:27)

While in excerpt (18) the long pause is probably meant to give time to the listener to observe the Arch before starting providing details about it, in example (19) the particularly long pause marks the end of a section focused on the development of the site, and the beginning of a new section focused on detailed description. The pause may also be length-related, since it occurs before an utterance made up of ten prosodic units.¹³

Finally, in *London by Rick Steves* pauses are generally over 1s, and few short pauses can be found between strictly interrelated contents, as in the following example:

- (20) By the way the word “Bride” in St. Bride’s is only coincidental. (0.5) The church was dedicated to St. Bridgit – or Bride – of Kildare long before the steeple or any wedding cakes. (1.2) St. Paul’s may be Wren’s most famous church, but smaller churches like St. Bride’s better illustrate his distinctive style. (London by Rick Steves, audio 2, min. 00:38)

As can be seen, the utterance occurring after the short pause is the explanation of the previous utterance, and then a longer pause occurs to mark the end of that subtopic.

To sum up, Table 4 reports the approximate extent to which pausing patterned in relation to content structure occurs in the three groups of audio fragments. The

¹³ In the wider research on prosody in audio guides (Fina 2016a), by ‘prosodic unit’ we mean a segment of speech occurring with a single pitch contour.

percentages were calculated taking into account both the number of audio fragments characterised by this pattern and the frequency of patterned pausing in each fragment:

ITA	UK	USA
20%	60%	30%

Table 4
Patterned pausing between utterances (% of cases).

These results will be discussed in the following section.

5. Discussion

The analysis of pausing between utterances in the selected fragments has shown that specific pausing strategies may be adopted in audio guides for effective content delivery and uptake. These strategies tend to be related to textual features, as they involve varying pause duration according to content organisation in topics and sub-topics. Looking specifically at the pausing strategies involved, particular attention was paid to pauses longer than 1s. It was noticed that these tend to occur in descriptive sections to focus the visitor's attention on a new item, especially when a shift in direction or location is implied. In a few cases, the long pause also gives time to the visitor to observe the new item. In narrative sections, long pauses tend to be used to signal the end of a subtopic and the beginning of a new one, while shorter pauses of the same duration tend to occur between utterances that are strictly interrelated in terms of content.

If we now refer to the key factors governing the distribution of pauses identified by Strangert (1991), we may argue that the one that governs pausing between utterances in audio guides is information load, as longer pauses are used to give time to the listener to absorb the previous information and cognitively prepare the visitor for new content. This pattern, though, tends to occur more systematically in the British audio fragments, whereas in the Italian and the American audio fragments long pauses between utterances tend to occur to a lesser extent, and the pausing strategies mentioned above tend to occur in a few audio fragments only. A tentative explanation to this divergence could be a different perception and understanding of silence across different cultures, with Italian and American audio guide producers probably tending to associate excessive silence between utterances to uncertainty, hesitation and distraction, as opposed to a presumably more positive attitude by British producers, according to which silence between utterances is an essential device for impressing information in the listener's mind. If true, this explanation would raise further cross-cultural issues, as similarities would be more plausible between the British and the American audio fragments – rather than between the Italian and the American ones – since they share the same language. Hence, further investigation on a wider corpus is needed to shed light on the reasons underlying this divergence.

In section 3, the dual stylistic nature of the audio guide genre led us to classify the audio guide as a sub-genre of public speaking, and more specifically as a particular type of script-based oral presentation, as the speaker introduces historical and artistic facts about the city in the same way as a human guide does, but reading aloud an elaborate script. We will now unfold the link between these two areas by framing the results of this study within best practices provided by manuals on public speaking. I decided to use as main reference framework the book by Cindy Griffin (2014), scholar and teacher of Speech

Communication at the Colorado State University, as it is a very recent and comprehensive publication.¹⁴ The book and the examples of public speaking it contains are entirely in English and, although no explicit reference to other languages is made, the book does not say that the guidelines provided are specifically for public presentations in English only. While they are certainly valid for any public presentation in English as a lingua franca (ELF), questions arise about whether they may be considered valid also for public speaking in Italian. A search on the Web showed that countless Italian books have been published on the matter. Among those consulted, the most comprehensive is that by Sansavini (2015), a communication trainer with a thirty-year experience in the field. This book covers the same aspects as Griffin's book (2014), and, most importantly, shares similar views and provides very similar guidelines.¹⁵

Griffin (2014, p. 312) defines pauses as part of 'timing', i.e. "the way you use pauses and delivery for maximum effect". She takes as example the use of pauses to deliver humour, and highlights the fact that "pauses to set a mood, before punch lines, or before key phrases can make the difference between a successful joke and one that falls flat". On a more general level, she advises against hesitation pauses (i.e. those filled with sounds like "er" or "um") in favour of intentional pauses, that should be used to punctuate words, establish mood, indicate a transition, or emphasise a point (2014, p. 227-229). Interesting points on the use of pauses are also raised by Bovée (2003, p. 299), who states that pauses can be inserted to "allow the audience a moment to think about an idea, to indicate a shift to a new idea or to a new section of your speech, or to heighten audience anticipation to your next idea".

If we now relate these aspects to the specific case of the audio fragments analysed in this study, we may argue that dosing pause duration can play a crucial role in content uptake. For example, too long pauses between strictly interrelated concepts may determine dispersion of information in the visitor's mind, in the same way as too short pauses between utterances focusing on different items of the same site, or treating different topics, may require higher cognitive effort for processing information. Thus, in terms of effective delivery, the diversification of pause duration in relation to content can be considered a good strategy in audio guides, as it helps the listener to process the information with little cognitive effort. At this point, two main intertwined issues are worth discussing: intermodal remediation and inter-lingual/cultural remediation. Intermodal remediation is relevant to the strict connections between pause position, pause duration and textual features identified in the analysis. The instances of patterned pausing identified mainly in the British audio fragments and in a few Italian and American audio fragments push the dual stylistic nature of the audio guide genre well beyond the theoretical level, towards a real need to prosodically make sense of the written script so as to remediate written information into intelligible, effective speech. This implies making a sort of 'musical arrangement' out of the script, by giving meaning to the various segments of information through a mindful distribution of pauses according to textual patterns. On the other hand, inter-lingual/cultural remediation comes into play when translating audio guides into

¹⁴ A 5th edition was published in 2015, but the 4th edition (2014) was more easily retrieved and used as a reference. A comparison of the table of contents of both editions showed that the contents are the same. As stated by the author, in the new edition further examples and readings have been added.

¹⁵ While examining other books written by Italian authors, it was observed that there is no variation in best practices between Italian and English, and this indicates that Italian best practices draw on the Anglo-American approach to public speaking. It is not by chance that the expression 'public speaking', searched on Google in the Italian domain, provides 125,000 results, and this further indicates that the Italians look at the Anglo-American best practices for public speaking.

different languages. In this case, questions arise about whether intermodal remediation should either stick to the prosodic patterns identified in the audio guide genre in that language, or whether intermodal remediation should rather be revised in translation in the light of the public speaking guidelines illustrated above, as long as these can be considered cross-culturally valid. In the former case, creating a global-English version of an Italian audio guide would be particularly challenging given the net differences in pausing patterns between the British and the American audio fragments. Sticking to the public speaking model for the prosodic remediation of the scripts certainly sounds as a more viable solution, especially when translating into English.

It goes without saying that the research presented in this paper can only be considered a pilot study for future, in-depth investigation of prosody in audio guides, which would be needed to shed light on the quite surprising similarities between the Italian and the American audio fragments, against any expectations on similar patterns between the British and the American audio fragments, as well as to test the expected effectiveness of the public speaking model on real audio guide users. Although the analysis was carried out on a limited number of city audio guides, the approach adopted by Strangert (1991), which combines prosodic, syntactic and textual perspectives¹⁶ to study read-aloud texts, could be considered equally valid for the investigation of pausing strategies in audio guides. This study also opens up a variety of possible perspectives for future research, which may involve studies on audio guide users to investigate their own perception of pauses and actual effects that pausing strategies may determine in content processing and uptake, as well as studies on audio guide speakers to investigate the extent to which the pausing styles they adopt are conventional rather than arbitrary, or mindfully enacted – rather than naturally produced – to determine specific effects.

In terms of implications for professionals operating in the field of tourist communication, this study highlights the need to tackle effectiveness in tourist oral communication by adopting a cross-disciplinary approach that include public speaking, in which the final product should be the result of different types of expertise melting together rather than of individual work.

Bionote: Maria Elisa Fina received her PhD in Translation and Intercultural Studies from the University of Salento (Lecce) in 2016, with a thesis on cross-language multimodal analysis of city audio guides in Italian and in English entitled *On effective audio guiding. A multimodal investigation of Italian, British and American audio guides*. Her research interests focus mainly on Italian vs English contrastive analysis in tourist communication mediated by new technologies, and also include travel-user generated content, Corpus Linguistics, and literary translation.

Author's address: elisa.fina@unisalento.it

Acknowledgements: This study has been possible thanks to the methodological and technical support by the *Centro di Ricerca Interdisciplinare sul Linguaggio* (CRIL) of the University of Salento directed by Prof. Mirko Grimaldi, in the persons of Prof. Barbara Gili Fivela and *Ing.* Francesco Sigona.

¹⁶The same approach has also been adopted to investigate strategies in utterance-internal pausing (Fina 2016a).

References

- Alfaro I., Nardon M., Pianesi F., Stock O., Zancanaro M. 2005, *Using cinematic techniques on mobile devices for cultural tourism*, in “Information Technology & Tourism” 7, pp. 61-71.
- Bové L.C. 2003, *Contemporary Public Speaking*, 2nd Ed. San Diego (California), Collegiate Press.
- Cappelli G. 2008a, *The translation of tourism-related websites and localization: problems and perspectives*, in Baicchi A. (ed.), *Voices of Translation: RILA Rassegna Italiana di Linguistica Applicata*, Bulzoni Editore, Rome, pp. 97-115.
- Cappelli G. 2008b, *Expats’ talk: Humour and irony in an expatriate’s travel blog*, in “Textus: English Studies in Italy” 21 [1], pp. 9-26.
- Cappelli G. 2012, *Travelling in space: Spatial representation in English and Italian tourism discourse*, in “Textus: English Studies in Italy” 25 [1], pp. 19-36.
- Cappelli G. 2013, *A perfect Tuscan experience’: Destination image and Cultural expectations in positive travel reviews in English and Italian*, in Bruti S. (ed), *Investigating the language-culture interface: English vis-à-vis Italian*, RILA Rassegna Italiana di Linguistica Applicata, Bulzoni Editore, Rome, pp. 59-86.
- Crystal D. 1994, *Refining stylistic discourse categories*, in Melchers G. and Warren B. (eds.), *English linguistics in honour of Magnus Ljung*, Almqvist & Wiksell, Stockholm, pp. 35-46.
- Crystal D., Davy D. 1969, *Investigating English style*, Longman, London.
- D’Egidio A. 2013, *The tourist gaze in English, Italian and German travel articles about Puglia: A corpus-based study*, in “ICAME Journal” 38, 57-72.
- D’Egidio A. 2014, *The Language of Tourists in English and Italian Travel Blogs and Trip Reports: a Corpus-based Analysis*, in “Lingue Culture Mediazioni” 1 [1-2], pp. 145-161.
- De Meo A., Vitale M., Pellegrino, E. 2015, *Voce e apprendimento multimediale*, in Romano A., Rivoira M., Meandri I. (eds.), *Aspetti prosodici e testuali del raccontare: dalla letteratura orale al parlato dei media*, X Convegno ASIV, Edizioni dell’Orso, pp. 219-228.
- Denti O. 2012, *The Island of Sardinia from travel book to travel guides: The evolution of a genre*, in *Textus: English Studies in Italy* 25 [1], pp. 37-50.
- Denti O. 2015, *Gazing at Italy from the East: a multimodal analysis of Malaysian tourist blogs*, in “Lingue Culture Mediazioni” 2 [15], pp. 47-68.
- Fina M.E. 2011, *What A TripAdvisor Corpus Can Tell Us About Culture*, in “Cultus. The Journal of Intercultural Mediation and Communication”, pp. 59-80.
- Fina M.E. 2016a, *On Effective Audio Guiding: A Multimodal Investigation of Italian, British and American Audio Guides*. PhD thesis deposited at University of Salento (Lecce).
- Fina M.E. 2016b, *Navigation and circulation in city audio guides: a comparison between Italian and English*, in Katan D. and Spinzi C. (eds.), *Tourism Across Cultures. Accessibility in Tourist Communication*, special issue of “Cultus. The Journal of Intercultural Mediation and Communication”, 9 [2]. Terni: Iconesoft, pp. 106-136.
- Francesconi S. 2011a, *Images and writing in tourist brochures*, in “Journal of Tourism and Cultural Change” 9 [4], pp. 341-356.
- Francesconi S. 2011b, *New Zealand as ‘the youngest country on earth’: a multimodal analysis of a tourist video*, in “Textus: English Studies in Italy” 24 [2], pp. 323-340.
- Francesconi S. 2014, *Reading Tourism Texts: A Multimodal Analysis*, Channel View Publications, Bristol.
- Francesconi S. 2015, *‘Dreaming in Italy’: ibridazione e creatività in un digital diary video*, in Rocca Longo M. and Pennacchia M. (eds.), *Turismo creativo e identità culturale*, Roma TrE-Press, Rome, pp. 45-57.
- Gavioli L. 2015, *Negotiating territories of knowledge: on interpreting talk in guided tours*, in “The Interpreters’ Newsletter” 20, EUT Edizioni Università di Trieste, Trieste, pp. 73-86.
- Gebbensleben S., Dittmann J., Vielhauer C. 2006, *Multimodal Audio Guide for Museums and Exhibitions*, in “Proceedings of SPIE Electronic Imaging – Multimedia on Mobile Devices II” 6074. [http://www.witi.cs.uni-magdeburg.de/~vielhaue/jabreflib/\[GeDV2006a\].PDF](http://www.witi.cs.uni-magdeburg.de/~vielhaue/jabreflib/[GeDV2006a].PDF) (12.02.2016).
- Griffin C.L. 2014, *Invitation to Public Speaking*, 4th Ed, Cengage Learning and National Geographic.
- Katan D. 2012, *Translating the tourist gaze: from heritage and ‘culture’ to actual encounter*, in “Pasos: revista de turismo y patrimonio cultural” 10 [4] [special issue], pp. 83-95.
- Maci S. 2013, *Tourism Discourse: Professional, Promotional and Digital voices*, ECIG, Genova.
- Manca E. 2013, *Describing through the five senses. A contrastive socio-cultural and linguistic analysis of Italian and British tourist websites*, in Manca E. and Bianchi F. (eds.), *Tourism and tourist promotion around the world: a linguistic and socio-cultural perspective*, Salento University Publishing, Lecce, pp. 109-124.

- Manca E. 2016, *Official Tourist Websites and the Cultural Communication Grammar Model: analysing language, visuals and cultural features*, in “Cultus. The Journal of Intercultural Mediation and Communication” 9 [1], pp. 2-22.
- Nardi A. 2012, *Sprachliche Handlungen in Audio-Guide-Texten zur bildenden Kunst*. In D. Höhmann (ed.), *Tourismuskommunikation, Im Spannungsfeld von Sprach- und Kulturkontakt*, Peter Lang, Frankfurt am Main, pp. 141-160.
- Nigro M.G. 2012, *From words to keywords: The journey from general language to the language of tourism*, in “Textus: English Studies in Italy” 25 [1], pp. 105-120.
- Neves J. 2012, *Multi-sensory approaches to (audio) describing visual art*, in Agost R., Orero P., Di Giovanni E. (eds.), *Multidisciplinarity in audiovisual translation*, special issue of *MonTi* 4, pp. 277-294.
- Neves J. 2013, *Guias Eletrônicas em Contexto Museológico – uma reflexão crítica*, in “Ensino Em Revista” 20 [1], pp. 163-178.
- Petrelli D., Not E. 2005, *User-centred design of flexible hypermedia for a mobile guide: reflections on the hyperaudio experience*, in “User Modeling and User-Adapted Interaction” 15 [3-4], pp. 303-338.
- Peverati C. 2012, *La ricerca linguistica negli Studi sul Turismo. Un repertorio bibliografico*, in Agorni M. (ed). *Prospettive Linguistiche e Traduttologiche negli Studi sul Turismo*, Franco Angeli, Milan, pp. 115-149.
- Plastina A.F. 2012, *Tourism destination image: Distortion or promotion? An analysis of web-based promotional discourse about Calabria*, in “Textus: English Studies in Italy” 25 [1], pp. 121-136.
- Ravazzolo E. 2014, *La visite guidée du jardin ethnobotanique: la transformation des objets naturels en signes culturels et verbaux*, in Dufiet, J.P. (ed.), *L’objet d’art et de culture à la lumière de ses médiations*, Trento, Università degli Studi di Trento, pp. 85-109.
- Sabatini F. 1997, *Prove per l’italiano ‘trasmesso’ (e auspici di un parlato serio semplice)*, in *Gli italiani trasmessi: la radio. Atti del Convegno*, Accademia della Crusca, Florence, pp. 11-30.
- Sansavini C. 2015, *Parlare in pubblico. Farsi capire, farsi ascoltare, persuadere il pubblico*, Giunti, Florence.
- Strangert E. 1991, *Pausing in Texts Read Aloud*, in “Proceedings of the XIIth International Congress of Phonetic Sciences”, pp. 238-241.
- Strangert E. 1993, *Speaking style and pausing*, in “PHONUM” 2, pp. 121-137.
- Strangert E. 2005, *Prosody in public speech: analyses of a news announcement and a political interview*, in “INTERSPEECH”, pp. 3401-3404.
- Tempel M., ten Thije J.D. 2012, *The appreciation of cultural and linguistic adjustments in multilingual museum audio tours by international tourist*, in Maier R. and ten Thije J.D. (eds.), *Managing Cultural and Linguistic Diversity in Multiple Organisation Settings*, special issue of *Journal of Multilingual and Multicultural Development* 33 [7], pp. 643-664.

Other sources

The Survey of English Usage (Quirk Randolph. 1959), <http://www.ucl.ac.uk/english-usage/> (12.02.2016).
 PRAAT, <http://www.fon.hum.uva.nl/praat/> (12.02.2016).

Websites of the audio guides (all last accessed February 2015)

<http://www.cittametropolitana.bo.it/turismo/Engine/RAServePG.php>
<http://www.galatamuseodelmare.it/cms/mirador-11837.html>
<http://www.materacittanarrata.it/downloads.asp>
http://turismo.provincia.milano.it/turismo/audio_video/audioguide/index.html
<http://www.padovanet.it/informazione/i-padova-percorsi-turistici-audioguidati>
<http://www.leaudioguide.net/a-it/Perugia>
http://www.comune.torino.it/torinoplus/Portrait_nuovo/itinerari/
<http://visitbristol.co.uk/things-to-do/sightseeing-and-tours/audio-tours>
www.tourist-tracks.com
www.realhistorytalks.com
[edinburghtour.com/Uncle Bob](http://edinburghtour.com/Uncle_Bob)
www.freelondontours.co.uk
<http://www.footnotesaudiowalks.co.uk/walks.php?Northampton/6/>
www.portsmouth-guide.co.uk

<http://interactive.wttw.com/loop/take-loop-tour>

<http://www.rmaguides.com>

http://www.frommers.com/rss/frommers_walking_tours.xml

<http://www.littlerock.com/things-to-do/audiotour/>

<https://www.ricksteves.com/watch-read-listen/audio/audio-tours/london>

www.welcomewalksusa.com

<http://www.learnoutloud.com/Catalog/Travel/-/Tour-of-the-National-Mall-and-Monuments-in-Washington/24833>