

RESEARCH ARTICLE

Do not sound like an announcer. The emphasis strategy in commercials

Emma Rodero^{1,2}  | Robert F. Potter³ 

¹Media Psychology Lab, Department of Communication, Pompeu Fabra University (UPF), Barcelona, Spain

²UPF Barcelona School of Management, Barcelona, Spain

³The Media School, Indiana University, Bloomington, Indiana, USA

Correspondence

Emma Rodero, Pompeu Fabra University (UPF), Roc Boronat, 138. 08108, Barcelona, Spain.

Email: emma.rodero@upf.edu

Funding information

FP7 People: Marie-Curie Actions, Grant/Award Number: 328636

Abstract

Most consumers do not like the speech style that announcers use in marketing messages. The main reason is an exaggerated way of speaking with a strong emphasis. Indeed, announcers tend to stress too many words in commercials, producing an overly emphatic speech style, which sounds choppy and stuttering. This study analyzes how different strategies of emphasizing words in commercials affect the listeners' cognitive processing. Four different strategies (no emphasis, moderate-low, moderate-high, and over-emphasis) were applied to 16 commercials. Participants ($N = 52$) had their physiological response (heart rate and skin conductance) measured during ad presentation. Afterward they assessed the commercials' effectiveness and adequacy and performed an immediate recall and a recognition test. The results indicated that the commercials presenting a moderate emphasis (low and high) improved the listeners' cognitive processing compared to messages with no emphasis strategy and the standard industry practice of using over-emphasis.

KEYWORDS

advertising, cognitive processing, memory, psychophysiology

1 | INTRODUCTION

In marketing, voice is one of the most valuable assets (Wiener & Chartrand, 2014). Most of the brand and advertising strategies are based on the announcers' voice, and consumers listen to them everywhere: on radio, television, YouTube, smart speakers, or phone. Therefore, the advertisement's effectiveness depends on how announcers command their voice (Wiener & Chartrand, 2014). This is the reason why it is essential to pay attention to this style. But when someone listens to advertising messages, they can quickly realize that radio and television announcers employ a marked speech style, also used by politicians or lectures in public speeches. This over accented speech style is exemplified by emphasizing certain words through increased duration, projection of greater intensity, or pitch

alteration to a high tone (Nadeu & Hualde, 2012). De-la-Mota and Rodero (2010) described this style as emphatic, explaining that announcers excessively highlight too many words in the message. Wheatley (1949) characterized it as misused pitch patterns to emphasize as many parts of a sentence as possible. Price (2005) defined this technique as using an excessive pitch range and emphasis. Van Leeuwen (1984) considered that announcers use this strategy to make relevant all the message parts.

The problem with this over-accentuation strategy is that increasing the number of emphatic words, it creates a choppy and stuttered delivery due to the excessive segmentation of prosody groups (De-la-Mota & Rodero, 2010). For this reason, some authors and producers argue that audiences might not feel that strategy is the optimal manner to deliver information in advertising messages.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. *Psychology & Marketing* Published by Wiley Periodicals LLC

To investigate this suggestion, we conducted an exploratory online survey, specifically for this study, with 200 respondents (female: 115; mean age: 39). A very high percentage (78.9%) of the respondents reported that they did not like the speech style used by announcers in commercials, and 67% of the respondents said that the main reason was the exaggerated way of speaking employed by in most ads—that is, ads utilizing a strong emphasis. On a 7-point Likert scale, the participants assessed naturalness as the most valued characteristic ($M = 6.7$; $SD = 0.46$), which is the opposite of an over-accentuation strategy.

Even given this perceptual data, however, we were aware of no research studies conducted to provide empirical data about the effect of this peculiar and widespread speech style on the cognitive processing of advertising messages. Therefore, following our survey we designed an experiment to analyze how different strategies of emphasizing words in audio commercials affect consumers' self-perception of the message's effectiveness and adequacy. Furthermore, psychophysiological measures were employed to provide evidence of the impact of different word emphasis strategies on consumers' attention and arousal. Memory tests conducted after ad exposure served as an index of recognition of information and immediate recall.

This investigation is underpinned by the Limited Capacity Model of Motivated Mediated Message Processing, also known as LC4MP (Lang, 2000, 2006). The theory posits that our cognitive system has limited resources to process information conveyed in media messages. These cognitive resources are allocated via automatic and controlled processing mechanisms (Lang, 2000). Controlled processing is conscious and demands comparably more effort than automatic processing, which is not under the control of the subject but is most often driven by elements of the message and the environment. One of the mechanisms through which resources are automatically allocated is the orienting response or reflex (Bradley, 2009; Lang, 1990). The orienting reflex is the response to a new stimulus that results in an allocation of resources to process it cognitively. This orienting response is triggered by two kinds of stimuli: those that are new or novel in the environment and those that a person perceives as important information. Some studies framed by this theory have shown that several structural features, as novel signals, can automatically allocate resources for cognitive processing (e.g., Potter et al., 2008). Some of these features are the announcer's voice (Potter, 2000) and pitch range (Rodero et al., 2017). The verbal emphasis strategy analyzed in this study is conceptualized as introducing novel signals into the auditory stream through the overtly emphasized words in an attempt to grab the listener's attention. Some studies have shown that speakers emphasize words with high information content when they speak to distracted addressees (Rosa et al., 2015). Also, this emphasis has a crucial influence on semantic processing (Cutler et al., 1997). Therefore, this theory seems applicable to an investigation of whether the strategy of emphasizing some words can also be a structural feature to grab attention and improve the cognitive processing of the message content. The ultimate goal is to determine how many emphatic words are optimal to increase the likelihood of novelty induction. Such knowledge can provide valuable information to marketers and media producers.

1.2 | The emphasis strategy

An announcer can underline the importance of critical information by emphasizing the most important words of a message. This strategy can be one of the most salient in guiding the listener, as long as it achieves the Principle of Distinctive and Contrastive Coherence of Prosody (Rodero, 2015). In this principle, the distinctive function distributes information and segments speech (Cutler & Norris, 1988). The strategy consists of highlighting only the essential information (If you have *problems* in *Mathematics*, we can *help*), which ultimately can benefit comprehension (Aylett & Turk, 2004; Calhoun, 2010; Terken & Nootboom, 1987). Some authors have shown that listeners assume this strategy to be used by a speaker (Arnold, 2008; Prince, 1981). Speakers tend to accent new information while the accessory parts tend not to be stressed (Brown & Yule, 1983; Cruttenden, 1997; Sityaev, 2000). In fact, some authors have demonstrated that the difficulty in encoding and comprehending information is more significant when the accent is incorrectly placed (Bock & Mazzella, 1983; Dahan et al., 2002). Studies also conclude that the rhythm created by accent in phrases helps when memorizing a series of words (Reeves et al., 2000; Sturges & Martin, 1974).

The second aspect of Rodero's (2015) Principle of Distinctive and Contrastive Coherence of Prosody, the contrastive function, is produced by the acoustic contrast between emphatic and nonemphatic words (If you have *problems* in *Mathematics*, we can *help*). Here the acoustic differentiation between targeted and untargeted words may draw the listener's attention (Li & Ren, 2013). Therefore, emphatic words could act as a novel signal and a claim of attention (Cutler & Fodor, 1979; Perrone et al., 2010). As some studies have shown, the targeted elements receive more attention than the rest (Cutler et al., 1977; Kristensen et al., 2013; Tong et al., 2005). This strategy is consistent with the Attentional Bounce Hypothesis (Pitt & Samuel, 1990). In this hypothesis, accented words may act like perceptual markers during processing, and, therefore, listeners would "bounce" one at a time. This acoustic contrast accomplishes the Spark Orientation Effect (Rodero, 2019). This effect describes a speech style to draw the listeners' attention to the relevant information of a message by using the so-called orienting elements (sound effects, appeals to the listener, or music). Therefore, this study will test if this effect is also produced by emphasizing key-words in the message, which could act as claims of attention.

Far from attracting consumers, the problem with the emphatic style that announcers commonly use is that it may produce a negative perception, as this strategy does not follow the Principle of Distinctive and Contrastive Coherence of Prosody (Rodero, 2015). The contrast to distinguish the critical part (new) from the accessory part (given) is eliminated here, as most of the words are highlighted. The same phenomenon can occur with the opposite strategy: not emphasizing any word. When some words do not dominate, communication tends to be monotone. Therefore, the distinctive and contrastive functions do not accomplish either. Listeners do not have a guide to know when they have to pay more attention. This means that listeners should poorly accept this strategy. In contrast, a moderate use, emphasizing only the

message's keywords in a medium quantity, could have a better effect, as this strategy fulfills the distinctive and contrastive functions. Therefore, we can post the first hypothesis:

- H1: Ads with a moderate emphasis strategy will get a higher effectiveness and adequacy rating than the ads with a no-emphasis or overemphasis strategy. Also, if the perception of the overemphasis and the no-emphasis strategies are negative, it is reasonable to expect that attention to the overall commercials may be low. However, the reasons for the lack of resource allocation are different for the two extremes of the emphasis strategy spectrum. In the case of overemphasis, previous research has demonstrated that audio productions at the highest levels of auditory structural complexity lead to attentional disengagement (Potter & Choi, 2006). Therefore, we predict that, in response to ads produced with an overemphasis strategy, listeners would disengage due to a negative response to the vocal enunciations themselves. Regarding the no emphasis strategy, it is logical to think that eliminating the acoustic signals will produce both negative perceptions and boredom. This will result in disengagement of cognitive resources from processing and comparatively less arousal response to message content. By comparison, in the moderate vocal emphasis strategy—when the speaker highlights only the keywords in a message, the emphasis can act as an orienting feature driving the listener's attention and increasing the level of activation or arousal (Lang, 2006; Rodero, 2019). In consequence, we hypothesize that the moderate emphasis strategy will result in the most listener engagement operationalized through psychophysiological data collected for the duration of the ads:
- H2: Ads with a moderate emphasis strategy will achieve higher physiological arousal and greater attention—as indexed by heart rate—compared with ads with a no-emphasis or overemphasis strategy. Also, if messages produced with a moderate emphasis strategy, that predicted to lead to greater listener engagement, better performance of the memory sub-processes of encoding and retrieval are also expected (Lang, 2000). Rodero (2015) showed that radio news recognition improved if only the essential words were emphasized naturally. Therefore, according to these studies, the H3 is formulated as:
- H3: Ads with a moderate emphasis strategy will achieve a higher recognition and recall level than the ads with a no-emphasis or overemphasis strategy.

2 | METHODS

The experiment was a 4 (emphasis strategy) by 4 (order of presentation) mixed factorial design. Each strategy had four commercials. Emphasis strategies and audio commercials were within-subject factors. The first strategy was produced with no intentional emphasis on any particular word, no planned salient pronouncements (No Emphasis-NE), the second had five important words emphasized across the ad (Moderate-Low Emphasis-MLE), the third had ten

important words emphasized (Moderate-High Emphasis-MHE), and the last one was delivered with fifteen important underlined words (Over emphasis-OE). We used four orders of presentation using the 16 ad stimuli. Participants were distributed in the four orders; thus, the order was a between-subjects factor.

2.1 | Participants

The participants were 52 students and staff at an American university. 29 were females, and 23 were males (19–32 years old). An a priori statistical power analysis was performed to calculate our sample size. With an $\alpha = 0.05$, power = 0.80, and effect size of 0.15, the projected sample size needed was 55. We recruited initially 55 participants, but three participants had to be removed, due to artifacts in their heartbeat register. Nonetheless, 52 participants is a sample considered ample in psychophysiological studies with a within-subjects design (Potter & Bolls, 2012). All signed the informed consent and received \$30 gift cards for participating in this experiment.

2.2 | Materials and stimuli

Audio advertisements were selected from the Radio Mercury Awards, a contest recognizing the most creative radio commercials in the United States. Sixteen ads were selected from the pool of all winners from the years 2000–2003. The criteria to choose these specific ads were: one voice-over and simple syntactic and grammar structures. The commercials were transcribed, modifying the names of the brands/products to avoid an easy recall. We also attempted to control the duration of the commercials, with the average being 22 seconds.

These 16 commercials were divided into the four levels of the strategies factor. Each level contained the same type of commercials, similar declarative sentences, and grammar structures. To decide which words should be elongated in the commercials and thus receive prominence, twenty individuals similar in age and education to the eventual study participants were surveyed to select the keywords in each commercial. The subjects were informed that only essential words (nouns, adjectives, or verbs) should be selected as opposed to function words (pronouns, prepositions, etc.). This was due to the likelihood that essential words are more likely to receive verbal emphasis given the concept of nuclear accent (German et al., 2006). The nuclear accent is the most important emphasis in the phrase, perceived as the most prominent (Calhoun, 2010).

At maximum, subjects selected 30% of a script (15 words) for emphasis. Given this, we chose the final number of keywords for emphasis as 5, 10, and 15, corresponding with 10%, 20%, and 30% of identified keywords in the whole 22-s commercial. For each strategy, the relative frequency of the proposed words was computed, and the most frequent words were finally selected (the first 5 for MLE, the first 10 for MHE, and the first 15 for the OE).

The scripts from the 16 commercials were recorded under optimal acoustic conditions by a professional male announcer with

resonant voice quality and relatively low pitch ($M = 115$ Hz, $SD = 30$); these characteristics have been described as optimal for male announcers (Borkowska & Pawlowski, 2011; Jones et al., 2008). The scripts highlighted the target words—those that should be elongated.

Emphasis can be marked with pitch, duration, loudness, or voice quality (Cooper et al., 1985; Fry, 1955; Kochanski et al., 2005; Ladd, 2008; Lieberman, 1960; Sluijter & Van Heuven, 1996). A very perceptible way to emphasize prominence can be word duration (Kochanski et al., 2005; Selkirk, 2008). In fact, focus information is produced with a longer duration (Breen et al., 2010; Frazier et al., 2006). From an expressive perspective, Rodero (2015) showed that when a speaker highlights words by prolonging them, participants perceived speech as more natural than when the speaker emphasized them by increasing pitch. According to these conclusions, in this study, the essential words were signaled by elongating them instead of using a marked pitch accent. The announcer was trained in producing the specific emphasis required in the experiment by prolonging the words.

The resulting recordings were acoustically analyzed using PRAAT (Boersma & Weenink, 2020) to ensure they were prepared according to the four specific strategies. First, an auditory analysis was performed to check that each commercial contained the expected number of prominences. Second, to test that the prominences were realized similarly across conditions, all focus words' duration was analyzed. The analysis results showed no significant differences among the targeted words' durational realization in the three strategies that utilized emphasis. The mean duration of the syllables in the MLE messages (five emphatic words) was 279 ms; in the MHE messages (10 emphatic words) was 343 ms, and in the OE messages (15 emphatic words) was 338 ms. This meant that the three levels were differentiated only by the number of emphatic words (quantity), but not by the manner to highlight them.

The 16 commercials, divided into four levels, were pretested with participants ($N = 25$) of the same age as the study participants. The goal was to assess the speaker's voice, the comprehensibility of the messages, and especially if the ads with these manipulations sounded like actual ads. Participants had to listen to the messages and rate them in an online survey using semantic differential scales of vocal effectiveness (how natural, pleasant, clear, persuasive, and dynamic the voice is) and adequacy (how credible, correct, appropriate and comprehensible the voice is). These 5-point scales have been used in other published studies of audio production strategies (Rodero et al., 2017) and were the same used for the main experiment. The pretest results confirmed the expected responses. The effectiveness ($M = 3.56$, $SD = 0.95$) and the adequacy index ($M = 3.89$, $SD = 0.89$) were above the average (2.50). The degree of comprehension for all the ads was also optimal, and no significant differences were found among categories, $F(3,22) = 364.62$, $p = 0.678$.

2.3 | Dependent variables

The dependent variables used were divided into three parts, as in Rodero et al. (2017): (a) self-reported measures of effectiveness and adequacy; (b) psychophysiological measures of arousal via skin

conductance level (SCL) and attention via heart rate (HR); and (c) cognitive tests to measure recognition accuracy and immediate recall.

Two 5-point scales evaluated the participants' perception of the ads: effectiveness and adequacy. These were the same scales used in the pretest. The effectiveness index was formed by creating mean values of how participants assessed the ads on the following dimensions—natural, pleasant, clear, persuasive, and dynamic. The adequacy index was formed by creating means of the level of comprehensibility, correctness, credibility, and appropriateness.

In this study, arousal was conceived as the sympathetic nervous system response and indexed as skin conductance or electrodermal activity (EDA) (Dawson et al., 2007). To do so, two electrodes, connected to Biopac 150, were placed on the participants' nondominant hand. The average of the minute before the onset of each audio presentation was used as the baseline. Response scores were calculated by subtracting the baseline values from the response signal. The data were prepared for analysis by computing change scores for each ad's 22-s duration and then collapsing across messages in each level of the strategy factor.

The Electrocardiogram (ECG) was collected and converted to heart rate (HR) offline to measure attention. Lower beats per minute (BPM) to environmental stimuli is also a physiological measure of attention (Potter & Bolls, 2012). Heart rate was measured by placing two electrodes on the participants' forearms. The ground electrode was on the non-dominant forearm. As in skin conductance, the previous minute to each audio presentation was used as a baseline. Response scores were calculated by subtracting the baseline values from the response signal. For the analysis, interbeat intervals were inspected for artifacts and premature beats. Segments containing artifacts were excluded. We used the Kardia software to clean the data (Perakakis et al., 2010). Three participants were removed from the analysis due to computer malfunction during data collection. The data were prepared for analysis by computing change scores for each ad's 22-s duration and then collapsing across messages in each level of the strategy factor.

We measured recall immediately after listening to the commercials. Participants had 30 s to write on the computer as many words as they remembered from the ad. The encoding subprocess was measured using recognition memory data collected after all 16 ads had been heard and after the participants had performed a distraction task. We performed a forced-choice test to measure the recognition of information. Participants listened to 48 phrases extracted from the radio ads (three different sentences for each commercial) and told to respond "Yes" or "No" on the computer keyboard according to whether they had heard the exact sentence in the previous ad stimuli. Foils were created by modifying two words in sentences using synonyms (i.e., The Cute Cat Tour, at College Avenue, for 1 day only.; The Cute Cat Tour, at Indiana Avenue, for 2 days only). The recognition hypothesis was tested using the number of hits or targets correctly identified.

2.4 | Procedure

Participants reported to the experimental lab individually at a pre-arranged time. A research assistant welcomed them and provided an

informed consent statement. Once the consent form was signed, they were escorted into a separate laboratory room and seated in a comfortable chair one by one. The physiological data collection procedures were verbally described, and electrodes were placed in the participants' palm and arms. Next, participants listened to the 16 messages. After each ad, the participant assessed the effectiveness and adequacy and completed the immediate recall test. The electrodes were removed after all ads, and participants performed a distraction task consisting of looking at still images and rating them (Lang et al., 2011). Participants then took the recognition test. After this test, they were asked if they had any questions about the procedures. Once all questions were addressed, the participant was given information about their compensation and dismissed.

3 | RESULTS

The results were analyzed by applying a 4 (strategy) \times 4 (order of presentation) analysis of variance (MANOVA) to the following dependent variables: effectiveness and adequacy of the voice, recall, and recognition of information. Each strategy had four commercials. The psychophysiological data were analyzed by applying 4 (strategies) \times 4 (ads) \times 22 (time) repeated-measures ANOVA for the variables skin conductance (arousal) and heart rate (attention).

3.1 | Hypothesis 1

The first hypothesis established that the messages with a ME would get a higher effectiveness and adequacy level than the messages with a NE or with an OE. Table 1 shows the mean values of the effectiveness and adequacy scales.

For effectiveness, the results were significant for the strategy type, $F(3,49) = 7.62$, $p < 0.001$, partial $\eta^2 = 0.028$. The order of presentation was not significant, $F(3,49) = 2.22$, $p = 0.084$, partial $\eta^2 = 0.008$. The MHE obtained the highest self-reported perceptions of effectiveness, followed by the MLE, NE, and OE strategies. The post hoc tests showed significant differences between the MHE and MLE ads and the OE and NE ads.

Regarding the adequacy index, the MHE strategy also achieved the highest levels, followed by the MLE, the NE, and the OE strategies (see Table 1). The results were significant for the strategy type, $F(3,812) = 12.16$, $p < 0.001$, partial $\eta^2 = .044$. The order of presentation was not significant, $F(3,49) = 1.61$, $p = 0.185$, partial $\eta^2 = 0.006$. The post hoc analysis showed significant differences between the moderate (high and low) emphasis with the rest. Therefore, these data support the first hypothesis.

3.2 | Hypothesis 2

The second hypothesis suggested that commercials with ME would achieve higher arousal and attention level than ads with NE or with

TABLE 1 Means and SD of effectiveness and adequacy

Strategy	Effectiveness		Adequacy	
	Mean	SD	Mean	SD
No emphasis	2.90	0.92	3.08	0.85
Moderate-low emphasis	3.13	0.79	3.43	0.76
Moderate-high emphasis	3.18	0.80	3.60	0.86
Over emphasis	2.82	0.81	3.09	0.93

TABLE 2 Means of skin conductance and heart rate

Strategy	Arousal		Attention	
	Mean	SD	Mean	SD
No emphasis	5	0.63	90.79	0.65
Moderate-low emphasis	5.25	0.57	88.05	0.65
Moderate-high emphasis	5.40	0.62	87.06	0.66
Over emphasis	5.10	0.60	90.60	0.84

an OE. Table 2 shows the physiological indices of arousal and attention collapsed over time.

Regarding arousal, the analysis showed that the MHE strategy elicited the highest skin conductance, followed by MLE, OE, and NE strategies. The results were significant for strategy type, $F(3,49) = 3.45$, $p = 0.043$, partial $\eta^2 = 0.091$, time $F(3,49) = 10.36$, $p < 0.001$, partial $\eta^2 = 0.047$, and for the interaction between strategy type and time, $F(3,49) = 2.46$, $p = 0.023$, partial $\eta^2 = 0.029$. There were no significant differences for commercials, $F(15,37) = 1.385$, $p = 0.245$, partial $\eta^2 = 0.027$, or the interaction commercials by time, $F(15,37) = 0.009$, $p = 0.923$, partial $\eta^2 = 0.000$. The post hoc tests revealed significant differences between the moderate strategies and the OE and NE emphasis. Figure 1 illustrates the dynamic data for skin conductance.

Third, the results for analysis of heart rate showed significant effects for time, $F(3,49) = 52.42$, $p < 0.001$, partial $\eta^2 = 0.02$, and for strategy by time, $F(3,49) = 5.24$, $p < 0.026$, partial $\eta^2 = 0.026$. There were no significant differences for commercials, $F(15,37) = 0.265$, $p = 0.609$, partial $\eta^2 = 0.005$, or the interaction commercials by time, $F(15,37) = 0.009$, $p = 0.955$, partial $\eta^2 = 0.000$. The lowest HR was for the strategy MHE, followed by MLE, OE, and NE strategies. The post hoc tests indicated significant differences between the medium strategies and the rest. Figure 2 shows the dynamic data for heart rate.

The data confirm the second hypothesis.

3.3 | Hypothesis 3

The third hypothesis suggested that the messages with ME should achieve a higher recognition and recall level than the commercials without a strategy (NE) or that OE. Table 3 shows the mean values of the memory measures.

Regarding recognition accuracy, measured by the number of correct responses, the greatest level strategies were the moderate

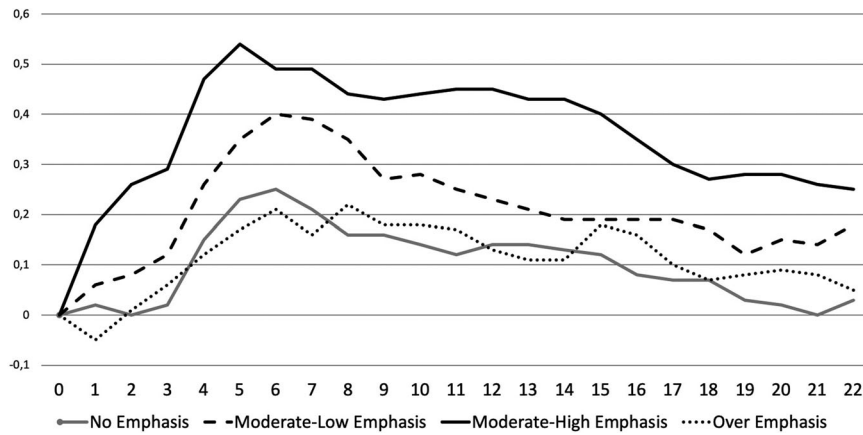


FIGURE 1 Skin conductance of the different emphasis strategies in 22-s commercials

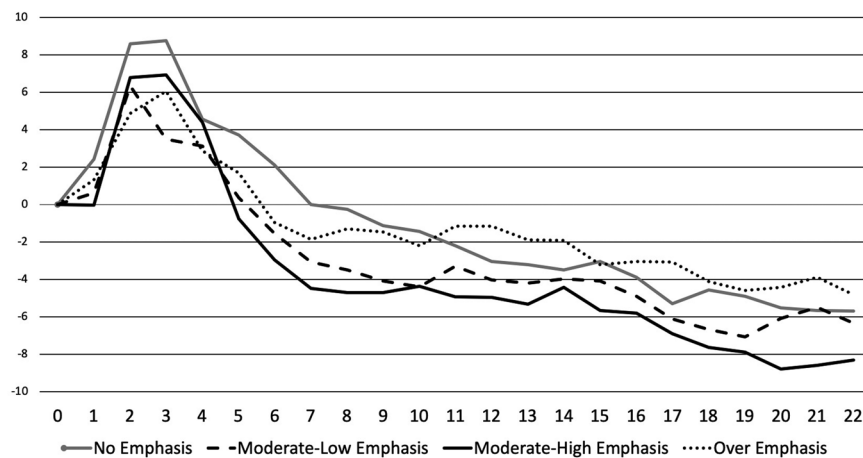


FIGURE 2 Heart rate of the different emphasis strategies in 22-s commercials

TABLE 3 Means and SD of recall and recognition

Strategy	Immediate recall		Recognition	
	Mean	SD	Mean	SD
No emphasis	19.70	6.88	0.63	0.024
Moderate-low emphasis	21.66	7.85	0.77	0.026
Moderate-high emphasis	21.18	7.94	0.76	0.022
Over emphasis	18.67	6.53	0.68	0.021

emphasis ads: MLE, followed by MHE. Then the extreme strategy commercials were in the lowest levels of recognition: NE and OE. The results were significant for strategy, $F(3,49) = 6.79$, $p < 0.001$, partial $\eta^2 = 0.069$. The order of presentation was not significant, $F(3,49) = 2.09$, $p = 0.099$, partial $\eta^2 = 0.003$. The post hoc analysis showed significant differences between the moderate emphasis (low and high) with the rest. This data supports the third hypothesis.

The immediate word recall was quantified with the percentage of correct recalled words, eliminating articles and prepositions, depending on the ad's total words. There were significant results for message strategy $F(3,49) = 7.14$, $p < 0.001$, partial $\eta^2 = 0.071$. The order of presentation was not significant, $F(3,49) = 3.56$, $p = 0.121$, partial $\eta^2 = 0.005$. The emphasis strategy with the greatest recall level was

the MLE, followed by MHE, OE, and NE. The post hoc analysis showed significant differences between messages created using moderate levels of emphasis (MLE and MHE) with the OE and NE commercials.

Of all recalled words (5, 10, or 15 depending on the ad), the emphatic percentage was 82% in MLE, 80% in MHE, and 60% in OE. Therefore, in moderate strategies, about 80% of the emphatic words were the most recalled of the commercial.

4 | DISCUSSION

This study aimed to determine whether the strategy of emphasizing keywords in commercials would influence self-perception of effectiveness and adequacy and the consumer's arousal, attention, recall, and recognition. The results showed that messages created using a moderate emphasis—with a medium number of elongated words—attained the best results in all the analyzed variables. This was particularly the case for the ads produced with MHE. In line with the LC4MP (Lang, 2006), all the results together revealed that the more activation of the sympathetic nervous system, shown by the results in skin conductance, combined with the increased cognitive resources allocated to encoding, indicated by heart rate data, improved recognition memory.

The results of this study showed that, regarding self-reported perceptions, the MHE strategy with ten emphatic words showed the best effectiveness and adequacy levels, followed by the MLE with five emphatic words, the NE without emphasis and, finally, the OE with 15 emphasized words. In the physiological response, the MHE obtained the best attention and arousal results, showing higher resource allocation and more sympathetic nervous activation. This strategy was followed by the other moderate strategy (MLE). The least position was for the ads using the two extreme strategies, the OE and the NE. We predicted this would be the case for the OE messages due to aversive reactions to the overemphasis strategy. In contrast, we thought the lack of enough verbally emphasized words would result in boredom for the NE messages. Finally, about memory, the superiority of a moderate strategy was also found. There were no statistically significant differences between the moderate emphasis (MLE and MHE), but the OE and the NE strategy commercials were significantly worse in words immediately recalled and subsequent delayed sentence recognition.

On the whole, these results showed the importance of a specific word emphasis strategy when attempting to influence the consumer's cognitive processing (Calhoun, 2010; Cevasco & Marmolejo, 2013; Cutler et al., 1997; Frazier et al., 2006; Kristensen et al., 2013). The strategy of emphasizing between 5 and 10 words in commercials improved effectiveness and adequacy, and enhanced cognitive processing, as these strategies had greater arousal, attention, and recognition levels than the other two ones (NE and OE). This strategy also helped the storage process, achieving the best levels of immediate recall.

Both linguistic and psychological theories may explain the underlying mechanisms of the moderate strategy's positive results. First, these commercials achieved the Principle of Distinctive and Contrastive Coherence of Prosody (Rodero, 2015). This principle emphasizes the information depending on their semantic meaning, highlighting the critical words in a sentence (Cutler & Norris, 1988). Second, in line with the Spark Orientation Effect (Rodero, 2019), the contrast caused by the words' emphasis discriminated this information acoustically (Kristensen et al., 2013; Tong et al., 2005), introducing a novel signal that contributed to grab the listener's attention (Cutler & Fodor, 1979; Perrone et al., 2010). This distinctive and contrastive effect applied to the message's keywords operated as a structural feature increasing attention, as shown in arousal and heart rate levels in moderate strategies with 5 and 10 emphasized words. The emphatic words might have reacted like a "flash" that conducted the participants' attention on the critical information. In consequence, the moderate strategies attained a positive perception. The number of emphatic words was convenient, as these strategies were assessed as effective and adequate. According to the LC4MP, listeners allocate more automatic resources when the message's structural features are simple, as in this case. Thanks to the simplicity of the message, participants had enough encoding resources. Therefore, the levels of recognition and recall were proper. In short, the emphasized information (contrastive) as the most important in the ads (distinctive) in an acceptable number (5/10 words

by commercial) increased both the acoustic and semantic contrast and, in consequence, the recall and recognition were improved.

These principles also may explain the negative results of the NE. When there is no strategy, the participants cannot determine the critical information to draw their attention. Here, both the distinctive and the contrastive functions were not accomplished, as there was no distribution of information according to its importance or novel acoustic signals to identify the key information. Therefore, the perception was not positive. The listeners said that this style was neither adequate nor effective. Furthermore, there were no attention claims without acoustic signals, and no resources were automatically allocated for encoding information. Then, the encoding process was affected, and the recognition of information was insufficient. Simultaneously, the storage process was hindered, and, thus, the recall was low. Finally, for participants was harder to process the ads with this strategy than the MLE and OE commercials.

However, the strategy with the worst results was the OE with 15 emphatic words. It is worth remembering that this style is the most used in advertising messages. However, as we explained previously, when the accents are placed incorrectly (Terken & Nooteboom, 1987), cognitive processing is more difficult (Bock & Mazzella, 1983; Dahan et al., 2002). This is the problem that this emphasis style could have had. This strategy only accomplished the distinctive function, but not the contrastive. The quantity of information per ad removed the contrast (if too much information is emphasized, contrast is reduced or even eliminated)—this problem made the cognitive processing hard. In this case, the Spark Orientation Effect (Rodero, 2019) and the attentional rebound (Pitt & Samuel, 1990) were eliminated. Thus, this strategy, commonly used in advertising, was assessed as few effective and adequate. The reason is that an overemphasis acoustically is perceived as a constant hammering, which sounds unnatural, and it is challenging to listen to for a long time. Along with perception, the arousal and attention levels were also altered. There was no acoustic contrast and, therefore, the differentiation between the key and not important was absent. This problem provoked that participants did not allocate automatic resources for encoding the information in the commercials, and there were no claims of attention. Lastly, both the storage and encoding processes were hindered, as shown in the recognition and recall levels. Cognitive processing was here superficial.

In brief, the most proper strategies for processing advertising messages, enhancing the perception (effectiveness and adequacy), the physiological response (attention and arousal), and the memory levels (recognition and recall) were commercials with a moderate level of emphatic words (5 and 10). The main conclusion is that the announcer has to emphasize only the critical information of the ad.

At a practical level, the study's findings provide valuable information to marketers and media producers: (a) by demonstrating that the current style used by announcers does not work and (b) by providing a strategy to improve the listener's cognitive processing.

First, in this study, we have shown that the overemphatic style is not the best strategy to convey commercials, as it causes problems in perception, attention, arousal, and memory. Therefore, announcers

should avoid using this common strategy in advertising. As it is very marked and characteristic, this speech emphasis shapes a very recognizable style that listeners have learned to associate clearly with advertising. The problem is that it is not an effective or adequate way to speak, as our study concludes, especially as it is very exaggerated, according to our previous survey results. This strategy is acoustically annoying, and for this reason, most respondents reported that they did not like it, mainly since it is exaggerated and not natural. Accordingly, when consumers listen to this typical emphatic way to speak, they automatically identify that this is a commercial, and then it is more likely that they decide not to pay attention and disconnect. Furthermore, this overemphasis eliminates the necessary prosody contrast to understand and follow the information. If too many keywords are emphasized in a message, the differentiation is eliminated, and it will be hard for listeners to know what the critical information is. For all these reasons, announcers should avoid this style.

Second, the main recommendation for advertising and marketing's announcers and voice-over artists is to use a moderate strategy. The moderate style contains enough words emphasized to attract attention by producing an acoustic contrast. As it differentiates the most important information of the message by emphasizing only the keywords, it can also improve memory. But at the same time that this style benefits attention and memory, as the emphasis is produced moderately, the style sounds more natural or unaffected, which was the most valued attribute for the participants of our previous survey. In conclusion, this moderate emphasis should be the strategy most used in advertising and marketing to optimize the results.

Along with advertising, this study and recommendations can be applied to many marketing strategies based on voice (voice marketing, voice branding, announcers in commercials, and phone calls), or audio (radio or podcast), or image (television or YouTube). Other areas can be management, business, political communication, public speaking, education activities, language learning, or speech synthesis.

This study can only be interpreted in the context of this lab experiment. One of the most important limitations and difficulties of this study was recording the analyzed styles modifying the conditions with a human voice, and not changing the speech signal with software after recording, as is usual in linguistic studies. Our goal was to obtain a natural speech produced by a speaker as close as possible to the real radio commercials. For this reason, we trained the professional speaker to produce the different conditions, and we had long recording sessions to get the expected result. Conversely, artificial manipulation has the advantage of controlling all the variables and achieving the exact condition in a more straightforward manner. However, we preferred having a more real commercial than a posterior voice manipulation that sounded artificial. Be that as it may, this is an aspect that should be taken into account. Related to this, although we controlled other prosody features, such as intonation or speech rate, it is clear that producing the stimulus with a speaker makes it more challenging to obtain a neutral speech that does not modify at all other speech elements. Anyhow, all the posterior analyses that we performed confirmed that the commercials in the same

condition sounded very similar. Along with this, another significant limitation is in the conditions that we selected (5, 10, and 15 emphasized words). It is evident that there are more possibilities and that the results can change depending on the number of marked words. However, this study can be understood as exploratory, as it analyzes a prosody element not examined previously. Future research can explore more options and, also, the combination with other prosody features. Finally, an important limitation is that in this study, we only examined a male voice. Further studies should study if there are differences when the emphasis strategy is conveyed by a female voice.

ACKNOWLEDGMENTS

This study was supported by a Marie Curie International Outgoing Fellowship within the 7th European Community Framework Programme (FP7-PEOPLE-2012-IOF, 328636). Thanks to the Institute for Communication Research in The Media School at Indiana University where this study was conducted.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Emma Rodero  <https://orcid.org/0000-0003-0948-3400>

Robert F. Potter  <https://orcid.org/0000-0002-9294-3940>

REFERENCES

- Arnold, J. E. (2008). Reference production: Production-internal and addressee-oriented processes. *Language and cognitive processes*, 23, 495–527.
- Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, 47, 31–56.
- Bock, J. K., & Mazzella, J. R. (1983). Intonational marking of given and new information: Some consequences for comprehension. *Memory and Cognition*, 11(1), 64–76.
- Boersma, P., & Weenink, D. (2020). "Praat: doing phonetics by computer (10.7)." <http://www.fon.hum.uva.nl/praat>
- Borkowska, B., & Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Animal Behaviour*, 82(1), 55–59.
- Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology*, 46(1), 1–11.
- Breen, M., Fedorenko, E., Wagner, M., & Gibson, E. (2010). Acoustic correlates of information structure. *Language and Cognitive Processes*, 25(7-9), 1044–1098.
- Brown, G., & Yule, G. (1983). *Teaching the spoken language* (Vol. 2). Cambridge University Press.
- Calhoun, S. (2010). How does informativeness affect prosodic prominence? *Language and Cognitive Processes*, 25, 1099–1140.
- Cevasco, J., & Marmolejo, F. (2013). The importance of studying prosody in the comprehension of spontaneous spoken discourse. *Revista Latinoamericana de Psicología*, 45, 21–33.
- Cooper, W. E., Eady, S. J., & Mueller, P. R. (1985). Acoustical aspects of contrastive stress in question-answer contexts. *The Journal of the Acoustical Society of America*, Jun 77(6), 2142–2156.

- Cruttenden, A. (1997). *Intonation*. Cambridge University Press.
- Cutler, A., & Fodor, J. A. (1979). Semantic focus and sentence comprehension. *Cognition*, 7(1), 49–59.
- Cutler, A., Fodor, J. A., & Foss, D. J. (1977). On the role of sentence stress in sentence processing. *Language and Speech*, 20, 1–10.
- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 113–121.
- Cutler, A., Norris, D., Dahan, D., & Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40, 141–201.
- Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, 47(2), 292–314.
- Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Bertsou (Eds.), *Handbook of psychophysiology* (3rd ed., pp. 159–181). Cambridge University Press.
- De-la-Mota, C., & Rodero, E. (2010). "La demarcación entonativa y el énfasis en la locución de los editores de boletines informativos radiofónicos," XXXIX Simposio Internacional de la Sociedad Española de Lingüística (SEL), Santiago de Compostela.
- Frazier, L., Carlson, K., & Clifton, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences*, 10, 244–249.
- Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America*, 27, 765–768.
- German, J., Pierrehumbert, J. B., & Kaufmann, S. (2006). Evidence for phonological constraints on nuclear accent placement. *Language*, 82(1), 151–168.
- Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., & Vukovic, J. (2008). Integrating cues of social interest and pitch in men's preferences for women's voices. *Biology Letters*, 4(2), 192–194.
- Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. S. (2005). Loudness predicts prominence: Fundamental frequency lends little. *Journal of the Acoustical Society of America*, 118, 1038–1054.
- Kristensen, L. B., Wang, L., Petersson, K. M., & Hagoort, P. (2013). The interface between language and attention: Prosodic focus marking recruits a general attention network in spoken language comprehension. *Cerebral Cortex*, 23, 1836–1848.
- Ladd, R. L. (2008). *Intonational phonology*. University Press.
- Lang, A. (1990). Involuntary attention and physiological arousal evoked by structural features and emotional content in TV commercials. *Communication Research*, 17(3), 275–299.
- Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication*, 50(1), 46–70.
- Lang, A. (2006). Using the limited capacity model of motivated mediated message processing to design effective cancer communication messages. *Journal of Communication*, 56, S57–S80.
- Lang, A., Kurita, S., Rubenking, B., & Potter, R. F. (2011). miniMAM: Validating a short version of the motivation activation measure. *Communication Methods & Measures*, 5(2), 146–162.
- Li, X.-qing, & Ren, G.-qin (2013). How and when accentuation influences temporally selective attention and subsequent semantic processing during online spoken language comprehension: an ERP study. *Neuropsychologia*, 51, 967–978.
- Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *The Journal of the Acoustical Society of America*, 32(4), 451–454.
- Nadeu, M., & Hualde, J. I. (2012). Acoustic correlates of emphatic stress in Central Catalan. *Language and Speech*, 55(4), 517–542.
- Perakakis, P., Joffily, M., Taylor, M., Guerra, P., & Vila, J. (2010). KARDIA: A Matlab software for the analysis of cardiac interbeat intervals. *Computer Methods and Programs in Biomedicine*, 98(1), 83–89.
- Perrone, M., Dohen, M., Loevenbruck, H., Sato, M., Pichat, C., Yvert, G., & Baciou, M. (2010). An fMRI study of the perception of contrastive prosodic focus in French. *Speech Prosody*, 100506, 1–4.
- Pitt, M. A., & Samuel, A. G. (1990). The use of rhythm in attending to speech. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 564–573.
- Potter, R. F. (2000). The effects of voice changes on orienting and immediate cognitive overload in radio listeners. *Media Psychology*, 2, 147–178.
- Potter, R. F., & Bolls, P. D. (2012). *Psychophysiological measurement and meaning: Cognitive and emotional processing of media*. Routledge.
- Potter, R. F., & Choi, J. (2006). The effects of auditory structural complexity on attitudes, attention, arousal & memory. *Media Psychology*, 8(4), 395–419.
- Potter, R. F., Lang, A., & Bolls, P. D. (2008). Identifying structural features of audio: Orienting responses during radio messages and their impact on recognition. *Journal of Media Psychology*, 20, 169–178.
- Price, J. (2005). "Allegations" and "Controversy": Are the Americans Invading our Intonational Space? *Conference of the Australian Linguistic Society*. Victoria, Australia: Australian Linguistic Society.
- Prince, E. (1981). Toward a taxonomy of given-new information. In E. P. Cole (Ed.), *Radical pragmatics* (pp. 223–255). Academic Press.
- Reeves, C., Schmauder, A. R., & Morris, R. K. (2000). Stress grouping improves performance on an immediate serial list recall task. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 1638–1654.
- Rodero, E. (2015). The principle of distinctive and contrastive coherence of prosody in radio news: An analysis of perception and recognition. *Journal of Nonverbal Behavior*, 39(1), 79–92.
- Rodero, E. (2019). The spark orientation effect for improving attention and recall. *Communication Research*, 46(7), 965–985.
- Rodero, E., Potter, R. F., & Prieto, P. (2017). Pitch range variations improve cognitive processing of audio messages. *Human Communication Research*, 43, 397–413.
- Rosa, E. C., Finch, K. H., Bergeson, M., & Arnold, J. E. (2015). The effects of addressee attention on prosodic prominence. *Language, Cognition and Neuroscience*, 30(1–2), 48–56.
- Selkirk, E. (2008). The relation between syntax and phonetics is indirect: the case of contrastive focus prosody. Experimental and theoretical advances in prosody. Cornell, NY, USA.
- Sityaev, D. (2000). The relationship between accentuation and information status of discourse referents: A corpus-based study, UCL Working Papers in Linguistics, 12.
- Sluijter, A., & Van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America*, 100, 2471–2485.
- Sturges, P., & Martin, J. G. (1974). Rhythmic structure in auditory pattern perception and immediate memory. *Journal of Experimental Psychology*, 102, 377–383.
- Terken, J., & Nöteboom, S. G. (1987). Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. *Language and Cognitive Processes*, 2, 145–163.
- Tong, Y., Gandour, J., Talavage, T., Wong, D., Dziedzic, M., Xu, Y., Li, X., & Lowe, M. (2005). Neural circuitry underlying sentence-level linguistic prosody. *NeuroImage*, 28(2), 417–428.
- Van Leeuwen, T. (1984). Impartial speech: Observations on the intonation of radio newsreader. *Australian Journal Cultural Studies*, 2, 84–98.
- Wheatley, K. E. (1949). Anomalies of radio speech. *American Speech*, 24(3), 213–215.
- Wiener, H. J., & Chartrand, T. L. (2014). The effect of voice quality on ad efficacy. *Psychology & Marketing*, 31(7), 509–517.

How to cite this article: Rodero, E., & Potter, R. F. (2021). Do not sound like an announcer. The emphasis strategy in commercials. *Psychol Mark*, 38, 1417–1425.
<https://doi.org/10.1002/mar.21525>