Auditioning the distinctiveness account:
Expanding the production effect to the auditory modality reveals the superiority of writing over vocalising

Yaniv Mama\textsuperscript{a} & Michal Icht\textsuperscript{b}

\textsuperscript{a} Department of Behavioral Sciences and Psychology, Ariel University, Ariel, Israel
\textsuperscript{b} Department of Communication Disorders, Ariel University, Ariel, Israel

Published online: 06 Dec 2014.

To cite this article: Yaniv Mama & Michal Icht (2014): Auditioning the distinctiveness account: Expanding the production effect to the auditory modality reveals the superiority of writing over vocalising, Memory, DOI: 10.1080/09658211.2014.986135

To link to this article: http://dx.doi.org/10.1080/09658211.2014.986135

PLEASE SCROLL DOWN FOR ARTICLE
Auditioning the distinctiveness account: Expanding the production effect to the auditory modality reveals the superiority of writing over vocalising

Yaniv Mama¹ and Michal Icht²

¹Department of Behavioral Sciences and Psychology, Ariel University, Ariel, Israel
²Department of Communication Disorders, Ariel University, Ariel, Israel

(Received 29 April 2014; accepted 4 November 2014)

The production effect (PE) documents the advantage in memory performance for words that are read aloud during study, rather than words that are read silently. Until now, the PE was examined in the visual modality, as the participants read the study words. In the present study, we extended the PE phenomenon and used the auditory modality at study. This novel methodology provides a critical test of the distinctiveness account. Accordingly, the participants heard the study words and learned them by vocal production (saying aloud) or by writing, followed by a free recall test. The use of the auditory modality yielded a memory advantage for words that were written during study over words that were vocally produced. We explain this result in light of the encoding distinctiveness account, suggesting that the PE is determined by the number of different encoding processes involved in learning, emphasising the essential role of active production.

Keywords: Memory; Production effect; Vocalisation; Distinctiveness; Learning modality.

In a standard production effect (PE) study, a list of study words is visually presented to the participants to be read and learned. Half of the list is read silently and the remaining half is read aloud. The typical result is enhanced memory for the vocally produced words (the PE). Forrin, MacLeod, and Ozubko (2012) expanded the basic PE definition, establishing that reading aloud holds an advantage over many other types of productions (e.g., reading a word and writing it, mouthing, typing or spelling). In other words, vocalising seems to be a superior mnemonic.

The leading theory of the PE to date is the encoding distinctiveness account. MacLeod and his associates (Forrin et al., 2012; MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010; Ozubko, Gopie, & MacLeod, 2012; Ozubko & MacLeod, 2010) suggest that the main mechanism underlying the PE is the execution of distinct, item-specific responses, vocal or non-vocal. According to this approach, the magnitude of the PE is determined by the amount of unique encoding processes involved in learning—the greater their number, the larger the memory enhancement. Therefore, the encoding distinctiveness account may predict the superiority of different productions, depending on the number of unique encoding processes.

Although not obvious at first glance, another look at the PE reveals that so far it has been treated as a modality-specific phenomenon, as the study words are always presented by the visual modality (reading). Is the PE restricted to visual...
presentation? Moreover, can this dependency explain the common finding of vocalisation superiority? The main goal of the present study was to expand the PE phenomenon by investigating it under a different study modality—audition. The modification of the study modality (auditory instead of visual) alters the number of distinct encoding processes involved in each of the different productions (vocalising and writing), thus offering a critical test of the distinctiveness account. The encoding distinctiveness account, which is a modal-independent theory, predicts a straightforward outcome: if the number of distinct encoding processes is the decisive factor, when study words are heard, writing will be a better mnemonic than vocalising.

The PE: Vocalisation improves memory

One can find evidence for the benefit of oral production on memory back in the 1970s of the last century. Crowder (1970) demonstrated the superiority of vocalisation (either by the participant or by the experimenter) over silent visual presentation. In 1972, Hopkins and Edwards had their participants study a list of words either by reading all of the words aloud or by reading all of them silently. Words studied aloud were recognised about 10% better than those silently read, but only in a within-subject condition (and not in a between-subjects condition). A pair of comprehensive studies by Conway and Gathercole (1987; Gathercole & Conway, 1988) thoroughly examined the effect of generating a word. In the first study (Conway & Gathercole, 1987), the authors compared the efficacy of five input conditions (types of production): (1) reading the word silently, (2) mouthing silently, (3) reading aloud, (4) listening and reading, and (5) listening. Consistent with Crowder (1970), the best memory performance was observed for words that were spoken aloud, regardless of the speaker (rememberer and experimenter). In the second study, Gathercole and Conway (1988) added writing (with or without seeing the written word) to the previous conditions and found that auditory presentation (the presence of acoustic information)—whether reading aloud, listening and reading, or merely listening—led to enhanced memory performance. Overall, Gathercole and Conway (1988) concluded that “acoustic presentation procedures led to better long-term recognition memory than non-acoustic presentation procedures” (p. 117). Furthermore, acoustic information improved memory especially when it was self-generated. The collective results of this pair of studies showed an advantage of 14–20% for reading a word aloud over the other methods tested (Conway & Gathercole, 1987; Gathercole & Conway, 1988).

In recent years, the vocalisation advantage described by Gathercole and Conway (1988) was rediscovered. MacLeod and his associates (2010) introduced this phenomenon, naming it “the production effect”. In a series of studies designed to investigate the boundaries of the PE, it was confirmed that a wide variety of productions improve memory. Spelling, writing, typing, mouthing and whispering—all enhance memory relative to silent reading. Yet, words that were read aloud were remembered better than all other means. Reviewing this study and others from the pertinent literature (Forrin et al., 2012; MacLeod et al., 2010), vocalisation stands out as the most pervasive means of bringing about a better memory. However, recently Quinlan and Taylor (2013) demonstrated that there are other forms of production that have a pronounced effect on memory than reading aloud. In their study, the participants were presented with words and instructed to either read them aloud loudly, read them aloud, read silently or sing (all are types of oral or vocal production). The authors report an advantage for reading items aloud loudly and for singing (the latter had a larger memory benefit than the former). They concluded that other forms of vocal production that include an additional distinct element produce greater memory benefits than reading aloud.

One should notice though, that in all of these experiments, the study words were learned visually. The words appeared on the computer screen and simultaneously were read (and produced) by the participants (Forrin et al., 2012; MacLeod, 2011; MacLeod et al., 2010; Ozubko et al., 2012; Quinlan & Taylor, 2013). In fact, the PE itself is defined as a phenomenon observed for visually presented stimuli. Is indeed the PE a modality-specific phenomenon? We addressed this issue in the present study, extending the PE to the auditory modality.

Distinctiveness as the source of the PE

Gathercole and Conway explained the advantage of vocalisation over other acoustic procedures by
the notion that self-generated cues enhance retrieval. One's own voice is more familiar, thus facilitating internal cue generation. In other words: "self-vocalized events may be comparatively effective because preexisting long-term knowledge structures are specialized for one's own voice" (Gathercole & Conway, 1988, p. 118). Along with this theoretical framework, Conway and Gathercole (1987) mentioned another important concept—distinctiveness. They suggested that acoustic attributes are highly distinctive, thus influencing memory. The enhanced memory for vocally produced words than for silently read words reflects a greater degree of general distinctiveness of the acoustic and articulatory actions than visual analysis of the silently read words.

The concept of distinctiveness was elaborated by MacLeod and his colleagues in order to explain the PE. According to this account, production at study (reading the word aloud) makes the sounded words stand out relative to the silently read words. This provides for a potent source of discrimination that may be used at test. Recollecting having said a word aloud during study can be used heuristically at test to guide memory decisions, enhancing memory performance. Forrin et al. (2012) concluded that the main mechanism underlying the PE is the execution of a distinct, item-specific response. The size of the PE is determined by the number of unique encoding processes involved. In this respect, the encoding distinctiveness account is not modal-dependent in nature, indifferent to study modality.

This simple reasoning explains why memory improves for words read aloud within visual presentation of the study words. Silent reading involves a single encoding process, whereas oral reading involves an extra pair of processes—(1) articulation (the execution of a motor action) and (2) audition (hearing oneself saying the word). Other learning conditions such as writing or mouthing also augment memory relative to silent reading because each entails one extra process of encoding (motor action). Since speaking aloud (vocalising) involves two additional distinct processes, it results in the largest memory advantage relative to all other methods of production.

Two pieces of evidence seem particularly conclusive in favour of the distinctiveness account. The first is the fact that in recall the PE is limited to within-subjects mixed-list designs in which only some of the words are vocalised during study (Jones & Pyc, 2014; MacLeod et al., 2010). Distinctiveness depends on relativity: pronounced words are salient on the background of other unpronounced words. Indeed, in pure-list conditions (when all the words are to be read aloud or silently), the PE is missing (Hopkins & Edwards, 1972). Similarly, the relative size of a subset of items to be produced (statistical distinctiveness) impacts on the magnitude of the PE (Icht, Mama, & Algom, 2014). The second piece of supporting evidence comes from the hierarchy of memory performance observed across silent reading, mouthing and reading aloud. Performance is perfectly correlated with the number of dedicated encoding processes: one in silent reading, two in mouthing and three in oral reading (Forrin et al., 2012; MacLeod et al., 2010).

Yet, it is important to note that in a recent meta-analysis by Fawcett (2013), a between-subject PE in recognition was demonstrated although smaller (and even non-significant within some individual studies) than the within-subjects PE. Bodner, Taikh, and Fawcett (2014) reported similar results (see also Dodson & Schacter, 2001, for a PE in false alarm rates observed between subjects as well). These findings under-mined the distinctiveness account and invited some alternative explanations of the PE. One of these explanations is the strength-based account (Bodner & Taikh, 2012; Bodner et al., 2014; Ozubko, Major, & MacLeod, 2010). According to this classic memory explanation (Wickelgren, 1969), reading words aloud strengthens their encoding and deepens their memory traces (representations) than reading words silently, regardless of the experimental design (within- or between-subjects) or list composition (mixed or pure). Since the strength account does not accord with the absence of PE on implicit-memory tests and its limited size in between-group designs, it seems that the distinctiveness account is still preferable (MacLeod et al., 2010; Ozubko & MacLeod, 2010; Ozubko, Major, & MacLeod, 2014).

Lastly, recent studies have suggested that in mixed lists, the PE may reflect a cost (decreased memory) to silently read words rather than a benefit (increased memory) for vocally produced words, in recognition (Bodner et al., 2014) and in free recall (Jones & Pyc, 2014). This cost may reflect superficial processing of silent words. Participants might use the silent trials as an opportunity to prepare for the next aloud trial, rather than focusing on encoding the silent words. Another possibility is that this cost occur because...
silent items are lazily read, receiving less attention than aloud items which require an overt response. Altogether, these results are consistent with the distinctive encoding account, as the costs to silent words can be attributed to relatively poor relational encoding (Hunt, 2013).

**Expanding the PE to the auditory modality**

Consider again the PE’s definition, as the benefit in memory for reading aloud relative to other distinct, item-specific responses (e.g., spelling, writing, typing and mouthing). What makes vocalisation so special? We suggest that the consistent finding of vocalisation advantage documented in the PE literature is due to the fact that all PE studies involved learning in the visual modality. Is it possible that using the auditory modality (listening to the study words) will yield different outcome?

In fact, back in the 1990s, Conway and Gathercole (1990) found that recognition of heard words was significantly enhanced by writing, relative to read words that benefitted less by this production. The authors explained these findings by translations between specialised processing domains (processing activities in different modalities). When such translations occur at encoding, they lead to the formation of distinctive memory traces and, hence, to better memory. According to this translation hypothesis, multi-modal encoding (that involves processes in different domains) is more robust than a single-modal encoding (involves only one modality).

In Conway and Gathercole’s (1990) words, “When ... an input activity requires translations between processing domains, then episodic memories are formed which are typically more distinctive than those arising from single-domain processing” (p. 515).

The translation hypothesis was questioned by a subsequent research by De Haan, Appels, Alemán, and Postma (2000) and was somewhat abandoned until recently. A new study by Rackie, Brandt, and Eysenck (2014) directly investigated the translation hypothesis and found translation effects (memory advantage for vocalising visually presented words and for writing auditorily presented words) in free recall and recognition.

Returning to the field of the PE, consider a typical PE experiment, with visual presentation of the study words. The participants are requested either to read the study words aloud or to write them down. Reading aloud (vocalising) involves three distinct encoding processes: (1) reading the word, (2) articulation (the execution of a motor action) and (3) hearing one’s own voice. Writing, on the other hand, involves only two distinct encoding processes: (1) reading the word and (2) writing (motor action). Note that following writing, the participants read again, their own handwrite this time, yet this last encoding process may not be fully distinct (as the first procedure was also reading). In fact, Forrin et al. (2012) used this rationale explaining the results of Experiment 2A, claiming: “writing involves one distinct motor process”, inferring that re-reading one’s own handwriting does not considered a distinct process. This simple logic explains the memory improvement for words read aloud over written words within visual learning.

Consider now broadening the PE methodology (and definition), using an auditory presentation of the study words. The participants hear the words (rather than read them) and requested either to say them aloud or to write them down. In this case, saying aloud (vocalising) involves two distinct encoding processes: (1) hearing the word and (2) articulation (the execution of a motor action). Although the participants also hear their own voice, this encoding process is not fully distinct (as the first procedure was also hearing). Production by writing, on the other hand, involves now three distinct encoding processes: (1) hearing the word, (2) writing (motor action) and (3) reading. According to the distinctiveness account, the additional encoding process involves in writing under the auditory presentation and predicts its superiority over vocalisation (Conway & Gathercole, 1990; Rackie et al., 2014). Thus, using the auditory modality changes the number of distinct encoding processes involved in each study condition, allowing a critical testing of the distinctiveness account, a modal-independent theory that has been derived from modality-specific methodology.

In the present study, we examine the effect of different learning modality, audition, on the PE. Specifically, our participants learned the study words by hearing them (and not by reading them) and produced them either by vocalising or by writing. In all other aspects, our methodology was similar to the standard PE experiments. Would learning the study material by the auditory modality (instead of by the visual modality) generate different results (e.g., writing > vocalising) from...
those typically obtained (vocalising > writing), suggesting that the PE is a modal-independent phenomenon (in harmony with its theoretical framework, the distinctiveness account)?

The present study

In three experiments, we tap the PE and its source by using an auditory presentation of the study words (instead of the common visual presentation). This systematic deployment of study modality may serve as a sensitive diagnostic tool providing a direct test of the encoding distinctiveness account. In Experiment 1, we seek to validate the presence of the PE under the auditory modality. In this initial experiment, the participants heard a list of words and learned them under three study conditions: (1) hearing (which involves a single encoding process), (2) hearing and silent reading (two different encoding processes) and (3) hearing and vocalising (three encoding processes, yet only two are different). Experiment 2 was similar to Experiment 1, except that the third study condition was hearing and writing (which entail three distinct encoding processes) instead of hearing and vocalising. Finally, in Experiment 3, we directly confronted these two types of productions—vocalisation and writing (within learning in the auditory modality), since typically the PE is limited to a within-subjects design (particularly in recall testing).

Note that in each of the experiments, at least one learning condition carries an active feature. Active encoding process refers to self-generated practice, an operative, explicit action one actually takes in order to improve memory and learning (Anderson, 2003). Referring to the active quality of production, Macleod (2011) concluded that “the production effect gains its value from the action of producing ... one’s own actions are more direct, more distinctive—more embodied—and hence more memorable due to their uniqueness” (p. 1201). In other words, as the action of production is more active, it is more distinct, thus enhancing memory to a greater extent (e.g., speaking and writing vs. hearing and reading).

By learning via the auditory modality, oral production carries three encoding processes, but only two are unique—hearing (the experimenter and one’s own voice) and articulating. Writing, on the other hand, holds an active feature (just as oral production), yet it carries now three different processes—hearing, writing and reading. This additional unique encoding process should result in a superior performance for writing than for vocalising in Experiments 2 and 3. By manipulating the learning conditions, we predict that the classic finding of oral production dominance can be reversed, and that writing will be most beneficial for memory following auditory learning.

EXPERIMENT 1

The main goal of Experiment 1 was to replicate and extend the classic PE finding of better memory performance for words said aloud, under an auditory presentation. The participants heard the study words and learned them under three different conditions: (1) “listen” (one encoding process), (2) “listen + read” (two distinct encoding processes) and (3) “listen + speak” (vocalising, which involved two unique encoding processes, one of which is active). Memory for study words was probed using a free recall test, in which studied words must be actively generated and retrieved by the participant, as it reliably reflects everyday memory abilities of retrieving without external cues (Castel, Rhodes, & Friedman, 2013; Conway & Gathercole, 1987 (Exp. 3); Forrin, Jonker, & MacLeod, 2014; Gathercole & Conway, 1988; Jones & Pye, 2014; Icht et al., 2014; Jonker, Levene, & MacLeod, 2014; Lin & MacLeod, 2012; MacLeod, 2011; Ozubko, Hourihan, & MacLeod, 2012 (fill-in-the-blank test); Putnam, Ozubko, MacLeod, & Roediger, 2014 (cued-recall]).

Method

Participants

A group of 22 young men and women, undergraduate students from Ariel University, received course credit for performing in the experiment. All were native Hebrew speakers and with hearing thresholds of at least 25 dB SPL for 500, 1000, 2000, and 4000 Hz (hearing screening was conducted using an AD229B inter-acoustic audiometer). Each participant was tested individually. The study was approved by the local ethics committee, and informed consent was obtained from the participants.

Apparatus and stimuli

The pool of items consisted of 80 Hebrew words, bi-syllabic nouns, three to five letters long,
with frequencies of greater than 12 per million (Frost & Plaut, 2005). The words were announced by a female speaker and recorded in a professional radio studio by Samplitude classic 8.1 program, using Sontronics TCS-6 microphone. From this pool, 45 words were selected for study, a different sample for each participant. During study, each of the 45 study words was auditorily presented via computer personal loudspeakers, at 35 dB SL (MCL), under the control of Direct-RT program.

At study, a random 15 of the words were simultaneously visually presented (“listen + read” condition). These written words appeared at the centre of a 15-inch colour monitor (Compaq laptop computer under control of Direct-RT program). The words were presented in black (28-point Arial), against a white background.

On each trial, a small icon (2 cm²) appeared in the upper part of the screen (approximately 5 cm above the centre). The icon entailed a small picture of an ear, of an eye or of a microphone. The icon indicated the appropriate learning condition for that word: the ear indicated the hearing condition (“listen”, 15 randomly selected words), the eye indicated hearing and silent reading (“listen + read”, another 15 words) and the microphone indicated hearing and vocal production (“listen + speak”, the remaining 15 words).

**Design and procedure**

**Study.** In each of the three conditions, the 45 study words were randomly divided into three subsets defined by the requested mode of learning. One subset (15 words) was learned by hearing, another subset (15 words) was learned by hearing and silent reading, and the other subset was learned by hearing and vocalising (the remaining 15 words). Each experimental trial started with a visual presentation of the icon (ear, eye or microphone). The study word was auditorily or auditorily and visually presented 300 msec following the icons’ appearance (the visual word appeared for 1 sec). After the item’s presentation, the icon remained visible for 3000 msec. A blank screen for 1000 msec followed (thus the interval between words was about 4 sec).

**Filler task.** A brief task of random number generation (see Miyake et al., 2000) followed the study phase, in order to reduce short-memory effects. The participants were instructed to randomly produce numbers between 1 and 9 during a time window of 1 min.

**Memory test.** Each participant was asked to write down from memory as many study words as she or he could recall. An empty sheet of paper and a pencil were provided by the experimenter.

The participants were tested individually in a dimly lit room. A research assistant was present in the room throughout the experimental session. Upon arrival, each participant read and signed the informed consent form. The participant was seated at a distance of 60 cm from the centre of the computer screen. The study words were played from two loudspeakers (Electro-Medical Instrument Co.), each positioned at the height of the listener’s head at 45° azimuth. The participant was told that the goal was to learn each word via the mode signalled by the icon (ear, eye and microphone) and that test of memory would follow the presentation of the words. The interval between words was 4 sec.

After the presentation of all the 45 words, the participants performed in a short (filler) task, generating numbers between 1 and 9 in a random fashion for 1 min. Free recall test followed, performed by writing down as many study words from memory as possible. The whole experimental session lasted approximately 20 min.

**Results**

Figure 1 gives the results of free recall (we calculated the number of correct words recalled, excluding intrusions). Plotted are the proportions of correct words recalled for the three learning conditions—“listen”, “listen + read” and “listen + speak”. Visual inspection reveals the superiority of vocal production over the other learning conditions (PE). Statistical analysis supports the conclusions based on the visual inspection of Figure 1. A repeated measures analysis of variance (ANOVA) with learning condition (listening, listening and reading silently, and listening and vocalising) as a within-subjects variable shows a significant effect $[F(2,42) = 8.918, p < .001, MSE = 2.418, \eta^2 = .298]$

Viewing Figure 1, it appears that vocalised words ($M = 33.6\%$) were better remembered than “listen” words ($M = 20.9\%$). A paired sample $t$-test confirms this difference $[t(21) = 4.28, p < .0005]$. Vocalised words were also better remembered than “listen + read” words ($M = 24.2\%$),
memory performance between the two conditions was absent in the former than the latter? Why was memory performance superior in the condition that involved two distinct encoding processes, only the “listen + speak” condition carries an active quality.

Note the similar memory performance in the “listen” and “listen + read” conditions. Despite the different number of distinct encoding processes (one vs. two), comparable recall rates were obtained. A possible explanation to the lack of gradient in memory performance is the notion that distinctiveness is relative (MacLeod et al., 2010). Hence, within a study list, a single subset of the words can stand out as distinct, while the remaining words serve as background, thus less distinct and less remembered. The same pattern of results was obtained by Gathercole and Conway (1988). Yet, these results can be attributed to low power due to the restricted number of participants (22 participants in the present experiment and 18 in Gathercole & Conway, 1988).

To recap, the results replicated the PE in an auditory modality, stressing that the PE is determined mainly by the presence of an active encoding process (saying aloud in this case). In Experiment 2, we further examined this assumption, by manipulating another active encoding process—writing. Previous studies have found this production to enhance memory (Forrin et al., 2012; but see: Gathercole & Conway, 1988) but less than vocalising under visual presentation. This finding was effortlessly explained by the encoding distinctiveness account. Writing a word involve two distinct processes, while vocalising entails three. However, in all these above-mentioned studies, the learning modality was visual, possibly determining the pattern of results. In the next experiment, by using the auditory modality (as was done in Experiment 1), we address this possibility, widening the theoretical framework of the distinctive encoding account (see Conway & Gathercole, 1990).

**EXPERIMENT 2**

The principle goal of Experiment 2 was to evaluate writing as another means of production. Three learning conditions were compared: “listen”, “listen + read” (both identical to those probed in Experiment 1) and a new condition—“listen + write”. Note that in contrast to former studies, we used (again) an auditory presentation condition. This may explain the better memory for the “listen + speak” condition. Although both “listen + read” and “listen + speak” conditions involve an equal number of distinct encoding processes, only the “listen + speak” condition carries an active quality.

Analysing the results in light of the encoding distinctiveness account (number of distinct encoding processes) is revealing. In the basic condition of “listen”, only one encoding process was involved. In the “listen + read” condition, a pair of distinct encoding processes can be identified. If the number of distinct encoding processes is indeed a crucial factor determining the PE then why these conditions did not differ in memory performance? Moreover, note that the “listen + speak” condition entails two distinct encoding processes, as in the “listen + read” condition. If so, why memory performance was superior in the former than the latter?

According to the encoding distinctiveness account, an active encoding process is a mandatory element underlies the PE. Clearly, this critical characteristic was absent in the “listen + read” condition. This may explain the better memory for the “listen + speak” condition. Although both “listen + read” and “listen + speak” conditions involve an equal number of distinct encoding processes, only the “listen + speak” condition carries an active quality.
of the study words. This departure from the typical experimental procedure carries significant consequence, as the number of distinct encoding processes in the “listen + write” condition was increased. Instead of two distinct encoding processes as in the former visual studies (reading the word and writing), the present methodology generates three distinct encoding processes in writing (listening, writing and reading the written word). Consequently, according to the distinctiveness account, it is the “listen + write” condition that should result in enhanced memory performance.

Method

Participants

A group of 21 young men and women, undergraduate students from Ariel University, received course credit for performing in the experiment. Inclusion conditions were identical to those of Experiment 1.

Apparatus and stimuli

These were the same as in Experiment 1, except that the microphone icons (indicating the “listen + speak” condition) that appeared with 15 of the study words in Experiment 1 were replaced with a pencil icon in the present experiment, indicating the writing (“listen + write”) condition.

Design and procedure

Study. In each of the three study conditions, the 45 words were randomly divided into three subsets defined by the mode of learning. One subset (15 words) was learned by hearing (“listen”), another subset (15 words) was learned by hearing and silent reading (“listen + read”) and the other subset was learned by hearing and writing (the remaining 15 words, “listen + write”). In order to allow the new writing condition, a pencil and a stack of notes were given to the participants at the beginning of the experiment. After writing down the appropriate study word (production), the note was taken away by the experimenter (in order to prevent further encoding of that word).

Filler task and memory test. These were the same as in Experiment 1.

All other aspects of the experimental design and procedure remained identical to Experiment 1.

Results

The mean proportion of correct words recalled for the three learning conditions—“listen”, “listen + read” and “listen + write”—is presented in Figure 2. The superiority of the “listen + write” condition over the other two learning conditions is clear. Statistical analysis confirms this observation, as a repeated measures ANOVA with learning condition (“listen”, “listen + read” and “listen + write”) as a within-subjects variable shows a significant effect \[F(2,40) = 9.58, \ p < .001, \ MSE = 2.073, \ \eta^2 = .324\].

A careful inspection of Figure 2 reveals that memory performance for the “listen + write” words (\(M = 26.6\%\)) was better than that of the “listen” words (\(M = 16.5\%\)), \(t(20) = 3.3, \ p < .005\) (paired sample \(t\)-test) and also than “listen + read” words (\(M = 14.6\%\)), \(t(20) = 3.54, \ p < .005\). Memory performance was not significantly different between the “listen” and the “listen + read” conditions, \(t(20) = .82, \ p > .2\).

We note a seemingly drop in memory performance between the “listen + read” conditions of Experiments 1 and 2 (24.2% and 14.6%, respectively). Yet, a between-subjects independent \(t\)-test yielded no significant difference.

Finally, in order to compare the two production conditions, the “listen + write” condition (Experiment 2) and the “listen + speak” condition (Experiment 1), a between-subject analysis was performed on standardised results (the actual number of recall words divided by the set average). This analysis yielded no significant difference (\(p > .5\)).

![Figure 2. Proportion of recalled words, calculated separately for the different study conditions. The error bars are standard errors of their respective means.](image-url)
Discussion

MacLeod and his colleagues defined the PE as “the finding that people have better explicit memory for words that they read aloud relative to words that they read silently” (Forrin et al., 2012, p. 1046; see also MacLeod, 2011; MacLeod et al., 2010). Nevertheless, they also suggested that other types of productions could improve memory, though in a lesser degree. This is the main finding of the present experiment, as learning by writing was superior to all other learning conditions (although the presentation modality was auditory).

We analyse the present results by the encoding distinctiveness account. In the basic “listen” condition, a single encoding process was involved. In the “listen + read” condition, two distinct encoding processes were present. In the “listen + write” condition, there were three distinct encoding processes. Obviously, the latter condition yielded better recall rates, as it carries the highest number of distinct encoding processes, as well as an active process of writing (as in Experiment 1, the two other subsets, “listen” and “listen + read”, did not differ from each other).

Reviewing the PE literature, vocalising has a large and consistent advantage over other types of production on memory. This was demonstrated in Experiment 1 under new learning modality—auditory. However, the results of Experiment 2 (again, using auditory presentation) revealed a memory advantage for writing (similar results were obtained by Rackie et al., 2014, but we noted that the study lists were much shorter). Comparing these two productions, the “listen + speak” condition of Experiment 1 and the “listen + write” condition of Experiment 2 (a between-subjects comparison) failed to show an advantage to neither production (F < 1). Consequently, a within-subjects comparison was called for (as in recall the PE is restricted to a within-subject design).

EXPERIMENT 3

Experiment 3 directly confronted two production types, the “listen + speak” condition of Experiment 1 and the “listen + write” condition of Experiment 2 in a within-subjects design. A similar confrontation was formerly performed by Gathercole and Conway (1988, Exp. 3), and by Forrin et al. (2012, Exp. 2A), but under the routinely used visual modality. Both experiments yielded superiority of vocal production over writing (and silent reading), the PE. Is it possible to alter this classic result (see Rackie et al., 2014)?

So far, two essential key variables of the PE were identified: (1) the execution of an active encoding process and (2) the number of distinct encoding processes. Regarding the first feature, both learning conditions, “listen + speak” and “listen + write”, involve an active processing. Concerning the second variable, “listen + speak” condition holds two distinct encoding processes: (1) hearing (the experimenter and oneself) and (2) vocalising. However, “listen + write” condition involve an extra distinct encoding process, as it carries (1) hearing the words, (2) writing them down (motor action) and (3) reading the written words, a total of three distinct encoding processes. Consequently, according to encoding distinctiveness, the “listen + write” condition carries higher relative distinctiveness. Therefore, the use of the auditory presentation modality may generate a surprising pattern of result—superiority for the “listen + write” condition over the “listen + speak” condition. In this respect, the present experiment aims to validate the theoretical basis of the PE, the encoding distinctiveness account, confirming its non-modal-dependent nature. If the PE is determined by the number of distinct encoding processes (regardless of study modality) then the “listen + write” condition should yield the best memory.

Method

Participants

A group of 18 young men and women, undergraduate students from Ariel University, received course credit for performing in the experiment. Inclusion conditions were identical to those of Experiment 1.

Apparatus and stimuli

A third of the study words (15) appeared with an ear icon, indicating the “listen” condition (15 randomly selected words), another third of the study words (15) appeared with a microphone icon, indicating the “listen + speak” condition. The remaining words (15) appeared with a pencil icon, indicating the writing (“listen + write”) condition.
Design and procedure

Study. In each of the three study conditions, the 45 words were randomly divided into three subsets defined by the mode of learning. One subset (15 words) was learned by hearing (“listen”), another subset (15 words) was learned by hearing and vocalising (“listen + speak”) and the other subset was learned by hearing and writing (the remaining 15 words, “listen + write”).

Filler task and memory test. These were the same as in Experiment 1.

All other aspects of the experimental design and procedure remained identical to Experiment 1.

Results

Figure 3 gives the proportion of correct words recalled for the three learning conditions—“listen”, “listen + speak” and “listen + write”. A quick glimpse of Figure 3 reveals an innovative discovery: a notable superiority of the “listen + write” condition over the other two learning conditions. Statistical analysis confirms this initial observation, as a repeated measures ANOVA with learning condition (“listen”, “listen + speak” and “listen + write”) as a within-subjects variable shows a significant effect \( F(2,34) = 11.447, p < .0001, \text{MSE} = 1.734, \eta^2 = .402 \).

A comprehensive examination of Figure 3 reveals that memory performance for the “listen + write” words (\( M = 27.7 \% \)) was better than that of the “listen” words (\( M = 14.4 \% \)), \( t(17) = 3.96, p = .0005 \) (paired sample \( t \)-test) and also than “listen + speak” words (\( M = 17.4 \% \)), \( t(17) = 3.91, p = .0005 \). The difference between memory performance for the “listen” and “listen + speak” conditions was insignificant \( t(17) = 1.09, p = .14 \).

Discussion

In the present experiment, we directly confronted two production types, “listen + speak” and “listen + write”, in a within-subjects design. The results revealed a reversal of the familiar pattern of results, as writing was found to be superior to vocalisation. In other words, participants remembered the study words under the “listen + write” condition better than words in the “listen + speak” condition. Previous attempts to oppose these two conditions (Gathercole & Conway, 1988, Exp. 3; Forrin et al., 2012, Exp. 2A) demonstrated memory enhancement for vocal production over writing (PE). So, what may be the source of these apparently opposing results?

The present experiment differs from previous PE studies solely by the modality in which the study words were presented—auditory vs. visual. This methodological modification carries an important and until now overlooked outcome: alteration of the number of distinct encoding processes. While in the usual visual presentation, writing entails two distinct encoding processes; in the auditory presentation, it carries three. On the other hand, vocalisation entails three distinct encoding processes in visual presentation, but only a pair in audition. Consequently, when the study words are heard (instead of seen), it is writing that improves their distinctiveness, rather than vocal production. This pattern of results is well fitted with the encoding distinctiveness account.

We note that the participants wrote words at study (under the “listen + write” condition) and later wrote them again at test (as the free recall test was a written one). Can this fact (the match between study and test modes) explain our results? A possible explanation can be found in the framework of transfer appropriate processing (Morris, Bransford, & Franks, 1977). This approach highlights that the value of specific encoding activities is determined relative to particular learning goals. Moreover, it stresses the relationship between learning processes and test requirements (the appropriateness of the testing situation). According to this approach, memory may be improved if the item processing required at study matches test demands (relative to situations in which the item encoding at study
mismatches test conditions). Thus, could the benefit for writing over speaking in Experiment 3 be due to participants writing the words at study and then writing the words again during test?

This possibility was ruled out by a recent study by Putnam and Roediger (2013). The authors tested whether first-test response mode (writing or speaking the answer) affects final-test performance and found no consistent advantage for one mode relative to the other. Moreover, final-test response mode itself did not matter, as the size of the testing effects was similar for both spoken and written responding on a first test when measured on a second test. In addition, final-test response mode did not interact with the first-test response mode.

At first glance, the results of Experiment 3 seem contradictory to those of Experiment 1, as the same learning condition, “listen + speak”, yielded better memory performance in Experiment 1 (34%) than in Experiment 3 (17.4%). Nevertheless, this difference may serve as a touchstone of the theoretical basis of the PE (distinctiveness vs. strength account). Recently, Bodner and Taikh (2012) suggested an alternative to the distinctiveness account, the strength account. According to this view, reading aloud increases the strength of the item’s memory trace relative to silent reading. Thus, stronger memory traces are formed for aloud words than silent words, which lead to better memory performance for the former. Other means can increase the strength of study words. Repeated presentations of silent words were used by Ozubko, Major, and MacLeod (2014). In fact, any type of production can enhance memory traces (e.g., writing the study words, as was done in the present Experiments 2 and 3). If indeed the decisive factor in the PE is the strengths of memory representations then vocal production would yield comparable memory performance in the “listen + speak” conditions in both Experiments 1 and 3, regardless of the competing learning conditions. Nonetheless, the difference between memory performance in the “listen + speak” conditions in Experiments 1 and 3 supports a (relative) distinctiveness account over a strength account. Since distinctiveness is relative, the production that involves the greater amount of distinct encoding processes makes items stand out at study. Thus, the “listen + speak” condition that was the most distinct is Experiment 1 was less distinct in Experiment 3 (relative to the “listen + write” condition). This alternation of distinctiveness yielded the change in memory performance.

Again, we suggested that two critical variables lie at the core of the PE, the execution of an active encoding process, alongside the number of distinct processes. As both learning conditions, “listen + speak” and “listen + write”, involve an active encoding, relative distinctiveness remains the main determinant. Relative distinctiveness is higher in the “listen + write” condition, thus making the words studied under this condition superior to the “listen + speak” condition.

**GENERAL DISCUSSION**

The PE is a well-known memory phenomenon refers to the benefit in memory for read aloud words relative to words read silently. Although other learning methods have been found to improve memory, it is vocal production that seems the robust and most consistent. A close examination of the pertinent PE studies reveals that they all used visual presentation modality. This fact arises an interesting pondering—is the PE a modality-specific phenomenon restricted to the visual modality? According to our findings, the answer is NO. Using an auditory presentation, in Experiments 1 and 2, we documented a memory advantage for vocally produced words (“listen + speak” condition) and for words produced by writing (“listen + write” condition) over non-produced words (“listen” and “listen + read” conditions). These results expand the boundaries of the PE to auditory learning. Experiment 3 directly compared the effect of vocalising vs. writing on memory under auditory learning, revealing the superiority of writing.

An early exploration of the effect of different types of productions (silent reading, mouthing, reading aloud and writing) on memory for written vs. heard words can be found in three studies by Conway and Gathercole (1987, 1990; Gathercole & Conway, 1988). The latter study is especially relevant to our study, demonstrating the efficiency of writing under auditory (relative to visual) presentation. These results have led to the translation hypothesis (Conway & Gathercole, 1990). According to this approach, input activities which require translations between input processing domains (such as reading and then saying a word that involve translations from orthography to phonology) yield distinct memory representations and, thus, to improved memory performance.
Two decades later, this memory phenomenon attracted renewed attention. In fact, two separate branches of research evolved from this early approach. The first remained very much faithful to the original translation hypothesis. For example, Rackie et al. (2014) examined the translation effects of vocalisation and writing in free recall and recognition. Other recent studies evaluated the translation effect of writing in note-taking in recall (Bui, Myerson, & Hale, 2013) and the translation effect of vocalisation on thinking aloud in intelligent problem-solving (Fox & Charness, 2010).

The second research field that emerged from the translation effect (and its spouse, the generation effect) is the PE (Forrin et al., 2012; Lin & MacLeod, 2012; MacLeod, 2011; MacLeod et al., 2010; Ozubko, Gopie, & MacLeod, 2012), up to now focusing on improving memory for written material. The prevailing theoretical explanation of the PE is the encoding distinctiveness account (MacLeod, 2011; MacLeod et al., 2010; Ozubko et al., 2012). This theory stresses the number of the different encoding processes as determines distinctiveness hence memory improvement. This theory is not modality-dependent in nature (just like the translation account). Accordingly, another purpose of the present study was to clarify this seemingly inconsistency between the PE modality-constrained methodology and the non-modal logic of the distinctiveness account.

Up till now, the encoding distinctiveness account easily explained the empirical results of the classic PE. Yet, an important variable eluded the common experimental design, as in the majority of studies vision served as the basic learning modality. This led to the common result of vocal production superiority. Using the auditory modality, writing was found to be the best mnemonic. According to the distinctive encoding approach, the larger the number of different encoding processes involved in learning, the greater the memory advantage. In the pertinent literature (visual modality), vocalisation showed a consistent advantage as it held the greater number of distinct processes. The results of the present study (auditory modality) are different, demonstrating the advantage of writing over vocalising. In three experiments, we employed a novel PE experimental design, using an auditory presentation of the study words. This shift from the traditional platform permitted an innovative perspective of the encoding distinctiveness account.

The novel use of an auditory modality at study changed the number of different encoding processes involve in each of the study conditions, hence their relative distinctiveness. When hearing a study word, repeating it aloud entails two different encoding processes. Although in this case, hearing occurs twice (first, hearing the recorded study word via speakers, and second, hearing oneself pronouncing it); both represent a single encoding process. As a result, in the present study, vocal production has been relegated and was replaced by writing as the best learning condition, since when hearing a study word, writing it down entails a trinity of different encoding processes. This straightforward logic easily explains the superiority of writing under auditory learning, supporting the distinctiveness account.

Interestingly, our present auditory PE study and the translation effect study by Rackie et al. (2014), deriving independently and simultaneously from different research directions, have used comparable experimental methodology (yet, note the difference in list length between the studies) and, more importantly, obtained similar results—superior free recall for writing auditory presented words. This pattern reinforces the findings and validates them.

The translation hypothesis requires translations between input processing domains. Hence, it necessarily involves greater number of distinct encoding processes relative to other (non-translated) productions. Different encoding activities (distinct modalities), such as reading a word (visual process) and then saying it aloud (auditory process), which require translations between input processing domains, enhance the distinctiveness of that word, hence its memory performance. Encoding activities that do not involve translations between modalities, such as reading a word (visual process) and then writing it (visual process again), enhance memory to a lesser degree. In this respect, the translation hypothesis and the encoding distinctiveness account share many features and have more in common than meets the eye.

Our pattern of results challenges the strength account that has been recently proposed as a theoretical basis for the PE (Bodner & Taikh, 2012; Bodner et al., 2014). According to this simple logic, aloud words are represented in memory more strongly than silent words. The strong memory traces of aloud words result in better memory performance (recall and
recognition) for such words than for silently read words. A strength-based account of the PE predicts equal memory performance for vocal production (best memory performance in the “listen + speak” condition), regardless of the other competing experimental conditions (“listen”, “listen + read” or “listen + write”). The different results of “listen + speak” conditions in Experiments 1 and 3 thus weaken the possibility that greater strength underlies the PE.

A careful reading of the literature regarding the strength account (e.g., Bodner & Taikh, 2012; Bodner, Taikh, & Fawcett, 2014; Fawcett, 2013; Fawcett, Quinlan, & Taylor, 2012) actually reveals a more integrated theory of the PE. Such model joined the strength-based account (a single process model, absolute in nature) as well as the distinctiveness account (a relative model), reducing the possible tension between them. Indeed, production may increase the strength of the relevant memory trace, and hence its distinctiveness, resulting in improved performance for produced items at test (Fawcett, 2013; Fawcett et al., 2012).

In contrast to this view, the distinctiveness account proposes that the act of production at study makes the produced words more distinctive, as they are encoded with extra detail. This distinctiveness can be used heuristically at test, as the participants can strategically access this information in order to better recall or recognise such words (MacLeod et al., 2010; Ozubko & MacLeod, 2010). Our results support that this additional information (greater amount of distinct encoding processes) underlies the PE.

According to this account, there are two variables defining the PE: (1) the execution of an active encoding process and (2) the number of distinct encoding processes (see Table 1, providing an overall look of the current experiments and summarising the relative weight of each variable). The same point was stressed by Forrin et al. (2012), stating that “speaking involves processing along an extra dimension relative to many other productions, processing that is more active and more embodied” (p. 1054).

The crucial role of an active encoding process arises from a careful inspection of Experiment 1. Words learned by the “listen + speak” condition were better remembered than words in “listen + read” condition, although both conditions hold similar number (two) of different encoding processes. The advantage of the former relative to the latter is attributed to the active process of speaking (articulating). On the other hand, reading is a receptive process and does not involve production. The extensive practice and long-term learning literate adults have had in reading (Dyer, 1973; Virzi & Egget, 1985) and have made reading seem virtually effortless and involuntary (i.e., automatic; Melara & Algom, 2003). This concept of automaticity (Hasher & Zacks, 1979; LaBerge & Samuels, 1974; Logan, 1980; Posner & Snyder, 1975) often explains Stroop effect (e.g., ignoring words when naming their print colours; Stroop, 1935). Consequently, in the context of PE, we consider the encoding process of silent reading as a passive one. Other encoding processes that are less automatic and more demanding (such as saying aloud or writing) are considered active procedures. Recently, Rackie et al. (2014) emphasised, “self-generated cues of writing and vocalisation acted as effective declarative memories … leading to the highest recall” (p. 7).

### Table 1

<table>
<thead>
<tr>
<th>Learning condition</th>
<th>Active encoding process</th>
<th>Number of different encoding processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“listen”</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>“listen + read”</td>
<td>–</td>
<td>2</td>
</tr>
<tr>
<td>“listen + speak”</td>
<td>+</td>
<td>2</td>
</tr>
<tr>
<td>Experiment 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“listen”</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>“listen + read”</td>
<td>–</td>
<td>2</td>
</tr>
<tr>
<td>“listen + write”</td>
<td>+</td>
<td>3</td>
</tr>
<tr>
<td>Experiment 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“listen”</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>“listen + speak”</td>
<td>+</td>
<td>2</td>
</tr>
<tr>
<td>“listen + write”</td>
<td>+</td>
<td>3</td>
</tr>
</tbody>
</table>
Expanding the PE to the auditory modality reveals that it is truly a modality-independent phenomenon, in congruence with the distinctiveness encoding theory. In addition to the theoretical importance of the present study, it carries practical implications. The current PE literature provides strong evidence supporting vocal production as the dominant method of memory enhancement. Yet, all these studies used visual learning. However, in everyday life, a great volume of the available information is auditory. Students in the classroom usually listen to their teacher, as frontal lectures (which involve oral presentation of the curriculum) are the most common teaching method of college instructors (Wirt et al., 2001). Talking to a friend over the telephone or listening to the radio—all involve acquiring information via the auditory channel. Even when watching the news on the television, the main portion of information is acoustic.

In line with former studies (Conway & Gathercole, 1990; Rackie et al., 2014), the present results suggest that production should be suitable to the situation (presentation modality). When reading a text (visual presentation), saying the words aloud is the best way to enhance memory. On the other hand, when listening to a lecture (auditory presentation), it is writing down that augments memory for the highest extent.

Not surprisingly, almost all college students take notes in class, in order to improve later memory performance (Dunkel & Davy, 1989; Palmatier & Bennett, 1974). Interestingly, the traditional taking notes by hand as well as the more modern version of typing them into a computer were both found to enhance memory (the latter strategy was even superior; Bui et al., 2013). In fact, individuals with poor working memory (who may have difficulties with writing notes) can successfully use such note-taking strategy of transcribing the lecture using a computer.

Caveats. Consider the latest findings regarding costs and benefits in the PE (Bodner et al., 2014; Bui et al., 2013). According to this view, in mixed lists, the PE reflects a cost (inferior memory) to silent items rather than a benefit (superior memory) for aloud items. Can production impair memory for non-produced items? One might consider the possibility the words in the “listen” condition, which is passive in nature (lazily heard?), are less encoded than words in the active conditions, “listen + speak” and “listen + write”. But, because pure-list groups were not tested in the present study, the impact of elaborative encoding on the benefits and costs of production could not be directly evaluated.

Our findings support the idea that distinctiveness is determined by the number of different encoding processes. Yet, the effect of the quality, uniqueness and extensiveness of each process is not clear, nor the way in which similar or overlapping processes interact in generating distinctiveness. A possible research direction may be a direct comparison of two types of productions, which involve similar amount of unique and active processes. For example, learning visually presented items via vocalising, and auditory presented material via writing. Both conditions include three separate encoding processes, one of which is active. Will these pairs of seemingly equivalent conditions equal memory performance? Future research is needed in order to further examine the importance of the number of distinct encoding processes involved in learning, as well as the active quality of encoding. Finally, the aforementioned similarities between the translation hypothesis and the encoding distinctiveness account call for further scrutiny that may lead to a unified new theoretical perspective.

REFERENCES


