Dickerson, W. (2015). A nail in the coffin of stress-timed rhythm. In J. Levis, R. Mohammed, M. Qian & Z. Zhou (Eds). *Proceedings of the 6th Pronunciation in Second Language Learning and Teaching Conference* (ISSN 2380-9566), Santa Barbara, CA (pp. 184-196). Ames, IA: Iowa State University.

A NAIL IN THE COFFIN OF STRESS-TIMED RHYTHM

Wayne B. Dickerson, University of Illinois at Urbana-Champaign

Stress-timed rhythm as applied to English (Pike, 1945) stands on three pillars. Research has convincingly invalidated two of these, undermining the claim that English is an exemplar of stress timing. If not stress timed, then what rhythm does English exhibit? This report describes a rhythm pattern that is not only widely attested in actual usage but is also simpler than the discredited pattern. Furthermore, learners are able to use it for clearer spontaneous speech.

INTRODUCTION

Prosody—rhythm and intonation—is a prime of spoken English; it gives the speech stream its sound shape (Bolinger, 1981). When the prosody of a phrase is right, a myriad of details of the language align to convey one's meaning without distraction; segmental errors can even pass unnoticed (Kjellin, 1999). When the prosody is wrong, the consequences for communication can be serious. This is why prosody is central to intelligibility (Derwing & Rossiter, 2003; Hahn, 2004). Recognizing this, many TESOL practitioners focus on developing learners' prosodic skills.

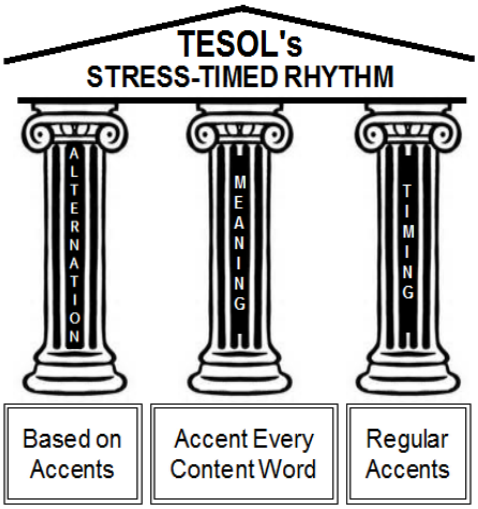
The questions we raise here are: In the area of rhythm, are we teaching the right content? If not, then what should we teach? Not long ago, these were pressing questions for this author, a TESOL practitioner and researcher. My reflections on a journey of loss and discovery are included.

THE CANON

Seventy years ago, when our field was young, insights moved from linguistics to ESL textbooks with little delay because most linguists were also language teachers. So when the linguist, Kenneth Pike, distinguished different types of rhythm, calling one "stress timed" and the other "syllable timed" in 1945, these distinctions spread quickly to the new profession (Prator, 1951; Lado & Fries, 1954).

Before long these ideas took root solidly in our TESOL narrative about the rhythm of English (stress-timed) and of most other languages (syllable-timed). They are now found in virtually every ESL pronunciation textbook, including mine (Dickerson, 1989/2004; Hahn & Dickerson, 1999a, 1999b). These rhythm types have the status of established fact, even though at the start they were untested hypotheses.

All of the features now associated with stress timing in TESOL were present in the 1950s: levels of stress—accented vs. unaccented,[1] the relative meaningfulness of English words—content words vs. function words,[2] and timing—a regular occurrence of accents. Although opinion was split at first about whether every content word should be stressed, the consensus came to support this view.[3] The results are the three pillars of stress-timed rhythm in the TESOL community.



THE TRAUMA OF LOSS

Although Pike uses stress timing and syllable timing to describe different rhythmic features of English, he also comments that syllable timing is dominant in Spanish (1945:35). Pushing this idea further, Abercrombie (1967) reasoned that, since the production of the two rhythm types requires mutually exclusive articulatory processes, languages should therefore exhibit one type of rhythm or another. His hypothesis explicitly addresses the third pillar above: accents come at regular intervals; it says nothing about the location of accents—the second pillar. His hypothesis challenged investigators to categorize languages according to their rhythmic pattern. It was disappointing work; for two decades researchers were unable to find a single instance of a stress-timed or syllable-timed language. Cauldwell (2002:1) summarizes:

---

[1] An accent is a heavy stress accompanied by a pitch change.

[2] Pike coined the terms "content word" and "function word." For him, content words are broader than "lexical words"—nouns, main verbs, adjectives, adverbs (of time, place, and manner). They also include demonstrative and indefinite pronouns, interrogatives, and interjections (Pike, 1945:118). All others are function words.

[3] The two dominant pronunciation textbooks at the time took different positions. Prator's 1951 text promotes stressing every content word; Lado & Fries (1954), following Pike's work, promotes deaccenting of content words. There are many reasons that Prator's position gained the ascendency besides the fact that his text stayed in print 20 years longer than Lado & Fries' text.

> The evidence from research is overwhelmingly against the hypothesis that languages are either 'stress-timed' or 'syllable-timed'. It is not possible to divide language into either 'syllable-timed' or 'stress-timed' categories; it is not the case that stresses occur at equal time-intervals in 'stress-timed' languages; it is not the case that syllables occur at equal time intervals in 'syllable-timed' languages; so-called 'syllable-timed' and 'stress-timed' languages are alike in having variations in syllable-length; so called 'syllable-timed' and 'stress-timed' languages are alike in having variations in inter-stress-interval length.[4]

My reaction to this conclusion was like that of many of my colleagues: Although the evidence was indeed overwhelming, it was still hard to drop stress and syllable timing. In the first place, the dichotomy offered a way to talk about differences I felt existed.[5] At first I wanted to keep talking about stress timing but with enough qualifiers to cover my bases—referring to "tendencies," "relatively regular pace," "stress-*based* rhythm" (Dauer, 1983)—all of which actually reject the claim of regularity in timing. In two contexts, however, this "fudge language" did not feel right: I did not want to continue telling an untruth to my MATESL students now that I knew better. Also in my ESL classes, I could no longer justify insisting that learners keep pace with a rhythm that was not really there. However, the biggest deterrent to letting go of stress timing was that I had no idea how to describe English rhythm if I did let go.

It took some time to realize that holding on to the myth of stress timing in any form would undercut any incentive to explore the reality of English rhythm. Ultimately, a determination emerged to find out what is going on with rhythm.

THE DELIGHT OF DISCOVERY

When I went in search of the real nature of English rhythm, I was surprised to discover that a number of scholars already had an answer. Even more surprising, they agreed on the kind of rhythm English has. Furthermore, they had been telling us about it for years. We TESOL professionals, however, were so confident about how English rhythm works that we had stopped paying attention. Three phonologists represent this group of researchers. They describe their conclusion about English rhythm using their own, but equivalent, terminology.

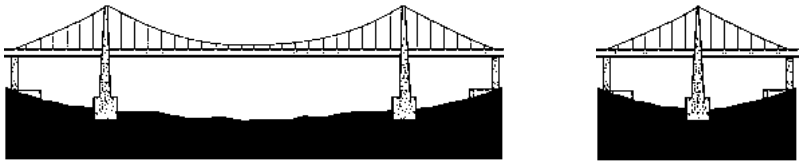
By 1961, the American linguist, Dwight Bolinger, had figured out the basic shapes of English rhythm. His metaphor of a suspension bridge represents a phrase, or what he calls a "melodic

---

[4]The evidence against stress timing: Instead of phrases being compressed or expanded to some uniform size, the length of inter-stress intervals is proportional to the number of syllables between the stresses (Roach, 1982; Dauer, 1983; Halliday, 1994; Bertrán, 1999). As early as 1939, Classé had tested the hypothesis, as articulated by Jones (1918), and found that the conditions necessary for strict stress timing to appear in everyday English speech are so rare as to occur only by accident (Cauldwell, 1996).

[5]We now know that perception and production are asymmetrical; up to a certain threshold of variability, we perceive more regularity in speech than is actually present (Lehiste, 1977, 1979). In fact, an irregular rhythm is functional; it keeps the listener attending to the message and not the tempo (Cauldwell, 2002).

group," with two accents—the two towers. He also describes a phrase with a single accent, but does not provide a graphic. If he had, it would look like the right-hand image below.



About this model of rhythm, Bolinger (1961) says,

> ...from the standpoint of running frequency they [melodic groups containing two accents and melodic groups containing one accent] probably outnumber all others in everyday conversational statements. (p. 135)

Bolinger is describing "everyday conversational statements"—the one- and two-second bursts of speech with which we all talk, commonly known as spontaneous speech.

Richard Cauldwell, a British scholar, also studies spontaneous speech. He cites frequency counts of "tone-units" in three studies and concludes (Cauldwell, 2002):

> The majority of speech can be divided up into tone-units of three sizes: single-prominence, double-prominence, and triple prominence tone-units... For all three [spoken] texts, single and double prominence tone-units account for very close to 90% of all [1700] tone units. (pp. 8, 15)

Phrases that have three or more accents are rare. They also tend not to be spontaneous speech. They represent instead "the successful delivery of a preplanned 'chunk' which may have been uttered before" (p. 17), snippets of verse, or "idiomatic, or semi-idiomatic material" (p. 18).

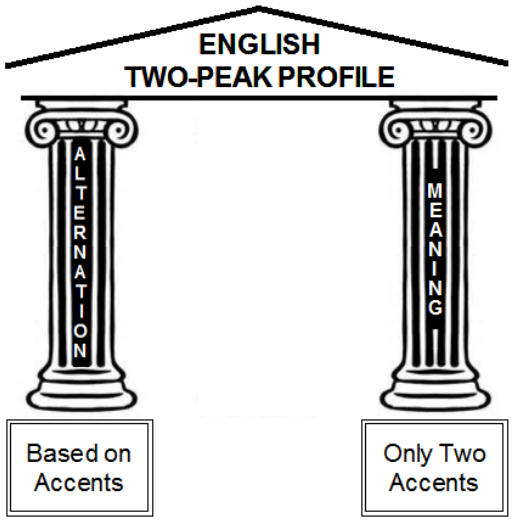
Another British researcher, John Wells, makes the same claim about naturally occurring 'intonational phrases' (IP) (Wells, 2006):

> An IP usually contains only one or two accents (onset and nucleus, or just a nucleus). (p. 192)

He notes that phrases with more than two accents are uncommon and are found more "in scripted material and in material read aloud" than in spontaneous conversation. Spontaneous phases are short and contain few accents because that small size is "the basic chunk for mental planning" (p. 192).

In short, whether referring to American English or British English, these scholars are telling us that a well-attested alternative to stress timing does exist. It is real, not hypothetical.[6] Like stress timing, the alternative is based on an alternation of what we call peaks (each with a single accent) and valleys (each with one or more lightly stressed or unstressed syllables); the first pillar (above) stands. Unlike stress timing, a spontaneous phrase has at most only two peaks according to the alternative model. If the phrase has more than two content words, then some of the content words will not carry a peak. Furthermore, none of these scholars considers timing to be at all relevant. These facts topple the second and third pillars of stress timing as applied to English.

It may seem strange to speak of rhythm without timing. However, this is what the research shows. Uniformity in the occurrence of peaks is not inherent in English (see footnote 4). Instead, rhythm is to be found in the repeated appearance of a pattern in the alternation, namely, two meaningful accents that surface in phrase after phrase. We call this the **two-peak profile**.



ACCESSING THE MODEL

If we accept that the two-peak profile is at play in everyday spoken English, then the practical question is: How do we introduce this new model to ESL learners?

Of the two peaks, we already know a lot about the accent called variously the nucleus, the tonic, the focal stress.[7] It is the single accent in one-peak phrases, and the second accent in two-peak phrases. In our ESL materials, it is the **primary stress** (●). By contrast, linguistic research tells us little about the first accent in two-peak phrases, generally known as the onset. In our teaching

[6] For spectrographic profiles of phrases with two accents see Dickerson, 2011, pp.75-76.

[7] Addressing ESL/EFL teachers, John Wells offers a generous, yet easily accessible, account of the nucleus (2006:93-186). A pedagogical counterpart for ESL/EFL learners can be found in Hahn & Dickerson (1999a:62-79; 1999b:80-81).

materials, we call it the **anchor** (○). If the learner cannot identify the anchor, it is impossible to create an authentic two-peak profile. So where is it?

At the University of Illinois we have been working on this question for several years. We have benefitted from having available transcripts from other linguists in which the anchor and primary stress have been marked on thousands of phrases in a discourse context. Together with our own data, we have pushed the analysis far enough to allow us to develop teaching materials for university-level learners. It has been encouraging to see that these learners understand. They can quickly make accurate predictions, and the rhythm they produce sounds like English rhythm. The second edition of our textbook, *Speechcraft,* is built around our research and what we have learned from trialing these materials with learners (Dickerson & Hahn, in press).

NAILING DOWN THE ANCHOR

Our anchor-placement strategy rests on five guiding positions—two goals and three procedures. First, our objective is to define the anchor in spontaneous speech, not in written texts delivered orally. The structure and length of phrases in the two styles of speech are radically different. The two-peak profile is native only to unrehearsed speech, the kind filled with short phrases limited by the speaker's and listener's natural on-the-fly processing powers.[8] Second, we aim to predict the anchor where it does not draw special attention to itself, in the neutral position. We discuss using the anchor for emphasis later. Third, we always start with the primary stress in place: Find the primary stress; then find the anchor. Fourth, we scan for the anchor at the beginning of the phrase and move to the right. Finally, we look for particular parts-of-speech.

Wells provides a preliminary hypothesis: "In unemphatic speech, only the *first* content word [left of the nucleus] receives an accent" (2006:207). Our test of this hypothesis includes these phrases.

○ ●

September's quite a full month.

○ ●

We're attending a conference in Santa Barbara.

○ ●

It's a really nice break for both of us.

[8] In our pedagogical materials, we call these phrases **message units**. They are equivalent to what Bolinger calls a "melodic group," what Cauldwell calls a "tone unit," and what Wells calls an "intonational phrase." In structure, they typically correspond to a grammatical unit and therefore express a coherent thought. In length, they are generally between one and seven words in length. Orally, they are most often identifiable by a pause and a phrase-final intonation pattern within which we find the primary stress. All of the following examples are instances of message units.

While only the first sentence conforms to the hypothesis, we still learn something important: The anchor likes **nouns**. It will hop over a verb and a variety of modifiers to reach a noun**.**

The phrases above also seem to show that the anchor does not like modifiers much. That generalization, however, is too broad, as the next phrases prove.

And we're really looking forward to it.

Maybe we'll even rent a car.

The anchor does not skip over all modifiers. If they modify a verb or a whole phrase, then the anchor seems to like **adverbs** fine.

Verbs, though, appear to be quite out-of-favor. The anchor hops over the verb in *We're attending a conference in Santa Barbara.* The following phrases refine the observation.

Take a look at the map again.

Let's visit Sequoia National Park.

For a verb to attract the anchor, it has to be either a second person imperative or a first person imperative. So **imperative** verbs attract the anchor.

To this point, the anchor seems to attach to some content words but not others. How does it like loud function words, namely, interrogative and negative words and demonstrative and indefinite subject pronouns?[9] The following sentences provide a clear answer.

Where would you like to camp?

He never packs a radio.

[9] In our pedagogical materials, words divide into three categories—content words, *loud* function words, and *soft* function words. Loud function words have the same potential for peak stress that content words have—nouns, adjectives, verbs, and adverbs. Soft function words usually carry valley stress—either tertiary or unstressed.

○ ●

That's the trail I was talking about.

○ ●

Did anyone remember the insect repellant?

Loud function words are solidly in the anchor camp.

On its first pass through a phrase, the anchor stops on a **N**oun, **A**dverb modifying a verb or whole phrase, **I**mperative verb, and any type of **L**oud function word, whichever comes first in a phrase. The acronym for this first pass is **NAIL**.

Of course, the first pass could yield no NAIL word left of the primary stress. In this case there are two possibilities. First, although there is no NAIL word, there may still be another content word as in these comments about giant sequoias.

○ ●

Hiking through a small grove of them, |

○ ●

we were truly speechless.

○ ●

They're as majestic as they are ancient.

The anchor lands on the first content word in the phrase: **1st CW**.

The second possibility is that there is no content word left of the primary stress. If not, the phrase has no anchor at all: **Ø**.

●

I can believe it!

●

You should have warned me.

With the basic framework in place, we turn to the last category of anchor-attracting words we call "number." It encompasses several related groups. The examples above demonstrate that the anchor hops over most modifiers on the first pass, particularly adjectives, but not all adjectives.

With only three days to sightsee, | [how far can we go?]

Fifty percent of the canyon | [is closed to tourists.]

The anchor clearly prefers a cardinal number if it encounters it first. Not surprisingly, the anchor reacts the same to ordinal numbers.

His first caution was about the bears.

Ordinal numbers are members of a larger group of "order" words, like *next, last, previous, following* as in sentences such as these:

Our last hike was to Half Dome.

The following dates are still open | [for a ranger tour.]

Order words also include superlatives—the *-est* form and adjectives and adverbs preceded by *most.* Superlatives identify the extreme end of some order or continuum.

The largest sequoia is the General Sherman.

The most challenging ascent for climbers | [is the 'Nose' of El Capitán.]

The anchor is attracted to either the single-word superlative or the adjective or adverb of the paraphrastic superlative, not the word *most.*

If we apply NAIL to these phrases with number and order words, we will mispredict the anchor on the first noun. To avoid this error, the anchor elects the first number or order word it encounters in the first pass. We include this category as "#": **#NAIL**.

The number/order category completes the sketch of our research. The anchor-placement formulation, **#NAIL or 1st CW or Ø**, is not the whole story, but it is enough to illustrate some profound consequences of the two-peak profile for learners.

EMPHATIC USES OF THE ANCHOR

The reason we insist on a neutral version of the anchor-placement rule is to show learners how to use the anchor for emphasis. An unusual placement of the anchor (or the primary stress)—that is, not in the expected neutral position—is understood as emphasis because this is how English speakers signal emphasis. Compare the neutral version of the sentences below with the following two increasingly emphatic versions. The first of these departs from the neutral rule by *moving* the anchor to a new position. The second, by *adding* a second anchor.[10]

	○ ●	○ ●
Neutral:	The evening was absolutely wonderful.	That's not what I meant.
	○ ●	○ ●
Emphatic1:	The evening was absolutely wonderful.	That's not what I meant.
	○ ○ ●	○ ○ ●
Emphatic2:	The evening was absolutely wonderful.	That's not what I meant.

When all content and loud function words are accented as prescribed by TESOL's stress-timed rhythm (the second pillar), not only is this rhythm unexpected, but it also sounds super-emphatic to the native ear. This is why native listeners are sometimes put off by the seeming forcefulness with which learners make their points. With the two-peak profile, learners speak in an unemphatic manner but understand the effect on the listener of moving or adding anchors.

CONCLUSION

There are good reasons to leave TESOL's version of stress-timed rhythm in the past and move on to accept the benefits of the two-peak profile as the foundation of learners' spontaneous speech.

1. Learners' rhythm will sound like *English* rhythm when learners use the *actual* rhythm of English as a model instead of an unnatural, artificial model. Specifically:
   - English speakers do not accent every content word and loud function word in unemphatic phrases.
   - The consensus of research is that accents do not come at equal intervals in spoken English.
   - An authentic alternative to stress timing is available in the two-peak profile.
2. Listeners expect to hear the two-peak profile and can decipher messages with an expected rhythm more easily than with an unnatural rhythm (Kjellin, 1999:11-12). Specifically:

[10] The fact that we can use the anchor emphatically proves there is a system to its placement. Otherwise anchor positions would not be noticed as emphatic. Furthermore, our ability to use the rule as stated here to identify emphatic usage suggests that our rule formulation, as simple as it is, captures reality.

- Stress timing generates peaks where valleys should be, thereby reducing the perceived prominence of the primary stress (Dickerson, 2011:75-78).
- Stress timing, when 'correctly' implemented, sends a strongly emphatic meaning in phrases containing multiple content and loud function words, a meaning which the learner is usually unaware of and does not intend.

As difficult as it is to say goodbye to stress timing, there is something quite powerful and authentic waiting for learners when we do let it go. We would like to think that this research may help put a NAIL in the coffin of stress timing.

ACKNOWLEDGEMENTS

The author wishes to thank Laura Hahn, Ester Cha, Yelena Tower, and Lonna Dickerson for their helpful feedback on this paper.

ABOUT THE AUTHOR

Wayne Dickerson is professor emeritus in the Department of Linguistics at the University of Illinois at Urbana-Champaign where he taught courses in English phonology (online and face-to-face) and ESL pronunciation. His research focuses on pedagogical applications of phonetics and phonology, pronunciation pedagogy, the value of orthography for learners, phonological variability, and pronunciation assessment. His two pronunciation textbooks are *Stress in the Speech Stream: The Rhythm of Spoken English* (1989), University of Illinois Press, and (with co-author Laura D. Hahn) *Speechcraft: Discourse Pronunciation for Advanced Learners* (1999), and associated workbooks, The University of Michigan Press.

Contact information:
Wayne Dickerson
7 Hale Haven Court
Savoy, IL 61874
dickrson@illinois.edu

REFERENCES

Abercrombie, D. (1967). *Elements of general phonetics.* Edinburgh: Edinburgh University Press.

Bertrán, A. (1999). Prosodic typology: On the dichotomy between stress-timed and syllable-timed languages. *Language design, 2,* 103-130.

Bolinger, D. (1961). Three analogies. *Hispanica, 44,* 135-136.

Bolinger, D. (1981). *Two kinds of vowels, two kinds of rhythm*. Bloomington, IN: Indiana University Linguistics Club.

Cauldwell, R. T. (1996). Stress-timing: Observations, beliefs, and evidence. *Eger journal of English studies, 1,* 33-48.

Cauldwell, R. T. (2002). The functional irrhythmicality of spontaneous speech: A discourse view of speech rhythms. *Apples, 2*, 1-24.

Classé, A. (1939). *The rhythm of English prose.* Oxford: Blackwell.

Dauer, R. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of phonetics, 11,* 51-62.

Derwing, T. M., & Rossiter, J. M. (2003). The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2 accented speech. *Applied language learning, 31,* 1-18.

Dickerson, W. B. (1989/2004). *Stress in the speech stream: The rhythm of spoken English.* Urbana, IL: University of Illinois Press. Published on CD in 2004 by the author.

Dickerson, W. B. (2011). Upstream destressing: Another step toward natural speech. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 2nd pronunciation in second language learning and teaching conference*, Ames, IA (pp. 70-81). Ames, IA: Iowa State University.

Dickerson, W. B., & Hahn, L. D. (in press). *Speechcraft: Discourse pronunciation for advanced learners.* (2nd ed.). Ann Arbor: The University of Michigan Press.

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL quarterly,* 38(2), 201-223.

Hahn, L. D., & Dickerson, W. B. (1999a). *Speechcraft: Discourse pronunciation for advanced learners.* Ann Arbor: The University of Michigan Press.

Hahn, L. D., & Dickerson, W. B. (1999b). *Speechcraft: Workbook for academic discourse.* Ann Arbor: The University of Michigan Press.

Halliday, M. (1994). *An introduction to functional grammar.* (2nd ed.). London: Edward Arnold.

Jones, D. (1918). *An outline of English Phonetics.* Leipzig and Berlin: B. G. Teubner.

Kjellin, O. (1999). Accent addition: Prosody and perception facilitates second language learning. In O. Fujimura, B.D. Joseph, & B. Palek (Eds.), *Proceedings of LP '98 (Linguistics and phonetics conference),* Ohio State University, Columbus, OH (pp. 373-398). Prague: The Karolinum Press.

Lado, R., & Fries, C. (1954). *English pronunciation.* The University of Michigan Press.

Lehiste, I. (1977). Isochrony reconsidered. *Journal of phonetics,* 5, 253-263.

Lehiste, I. (1979). The perception of duration within sequences of four intervals. *Journal of phonetics,* 7, 313-316.

Pike, K. L. (1945). *The intonation of American English.* Ann Arbor: The University of Michigan Press.

Pike, K. L. (1942). Pronunciation. In *Intensive course in English for Latin-American students*. Vol 1. Ann Arbor: The University of Michigan Press.

Prator, C. L., Jr. (1951). *Manual of American English pronunciation for adult foreign students.* Los Angeles: University of California Press.

Roach, P. (1982). On the distinction between ‘stress-timed’ and ‘syllable-timed’ languages. In D. Crystal (Ed.), *Linguistic controversies*, 73-79. London: Edward Arnold.

Wells, J. C. (2006). *English intonation: An introduction.* Cambridge: Cambridge University Press.