|  |  |  |  |
| --- | --- | --- | --- |
| **Instructions for ROCLING 2025 Proceedings (中文稿件需提供英文題目)** | | | |
|  | | | |
|  | | | |
|  | **First Author**  Affiliation / Address line 1  Affiliation / Address line 2  Affiliation / Address line 3  email@domain | **Second Author**  Affiliation / Address line 1  Affiliation / Address line 2  Affiliation / Address line 3  email@domain |  |
|  | | | |

摘要

發音品質評估中的發音好壞度 (Goodness of Pronunciation, GOP) 分數，是電腦輔助語言學習的關鍵技術 。近期的研究指出，直接使用聲學模型原始輸出logits 來計算 GOP 分數，其表現優於傳統基於 softmax 機率的方法，因為 logits 避免了機率飽和問題並保留了更豐富的區分性資訊 。然而，現有的logit-based 方法大多僅依賴最大值、均值或變異數等基本統計量 ，這忽略了在音素持續時間內，logit 序列更為複雜的動態分佈與時序特性。為了更全面地捕捉 logit 序列中所蘊含的發音細節，本研究提出了一套多面向的統計分析法。我們探索了五種能夠描述 logit 序列不同特性的高階統計指標：(1) 動差生成函數，用以計算分佈的偏度 (skewness) 與峰度 (kurtosis)；(2) 資訊理論，透過計算熵 (entropy) 來量化模型的不確定性；(3) 高斯混合模型 (GMM)，用以擬合 logit 的多模態分佈；(4) 時間序列分析，計算自相關係數 (autocorrelation) 來衡量 logit 的穩定性；以及 (5) 極值理論，採用 top-k 平均來獲得更穩健的峰值信心度估計。我們在公開的 L2 英語語音資料庫 (SpeechOcean762) 上進行實驗 ，將這些新提出的統計指標與參考文獻中的基線方法 (GOP\_MaxLogit, GOP\_Margin) 進行效能比較 。初步結果顯示，部分高階統計指標，特別是那些能夠描述logit 序列穩定性和分佈形狀的特徵，在發音錯誤檢測的分類任務上展現出更高的準確性，並與人類專家評分呈現出更強的相關性。這項研究證明，對 logit 序列進行更深層次的統計建模，是提升自動化發音評估系統效能的一個有效途徑。

Abstract.

關鍵字：關鍵字1、關鍵字2

Keywords: Keyword 1, Keyword 2

Introduction

在全球化時代，第二語言的口語溝通能力對於學術。然而，清晰的發音對L2學習者而言充滿挑戰，主要是因為母語 (L2) (L1) 的語音習慣會造成持續性的發音錯誤 。為此，電腦輔助發音訓練 (CAPT) 系統被廣泛發展，以提供即時且客觀的發音回饋 。在 CAPT 系統中，能夠在音素 (phoneme) 層級進行的發音錯誤檢測 (Mispronunciation Detection)，被證實對學習者改善特定發音問題特別有效 。

發音好壞度 (Goodness of Pronunciation, GOP) 是目前最主流的音素級別自動評估指標之一。傳統上，GOP 分數的計算依賴於深度神經網路 (DNN) 聲學模型輸出的後驗機率 (posterior probabilities) 。這些機率值是透過對模型的原始輸出logits 進行 softmax 歸一化得到的 。然而，softmax函數本身存在著「過度自信 (overconfidence)」的缺陷，容易將機率分佈推向極端，從而壓縮了不同音素之間的區分度，使得一些細微的發音偏差難以被偵測 。

為了解決 softmax 歸一化的限制，Parikh et al. (2025)的研究開創性地提出直接使用未經處理的logits 來計算 GOP 分數。相較於機率值，logits 保留了更豐富的鑑別資訊，並且避免了梯度飽和問題 。該研究探索了幾種基於logit 的指標，例如最大 Logit (GOP\_MaxLogit) ，用以捕捉模型的峰值信心；Logit 邊界 (GOP\_Margin) ，用以量化目標音素與其最主要競爭者之間的分離程度；以及Logit 變異數 (GOP\_VarLogit) ，用以衡量模型信心的穩定性。他們的實驗證明，在多數情況下，logit-based 的方法在發音錯誤檢測的分類任務上優於傳統的機率方法 。

儘管 Parikh et al. 的研究為 GOP 計算開闢了新的方向，但我們認為，他們所使用的方法仍有其侷限性。這些指標主要依賴 logit 序列的單點統計量 (如最大值) 或一階動差 (如均值、變異數)。這相當於將一個音素在持續時間內的 logit 變化視為一組無序的數字集合，忽略了其作為時間序列的內在結構以及其統計分佈的完整「形狀」。一個發音的過程是連續且動態的，其對應的 logit 序列在時間維度上的穩定性、對稱性與峰銳度，理應蘊含著關於發音品質的更深層線索。

基於此觀點，本研究旨在「超越均值與變異數」，提出一套更為全面且多面向的 logit 序列統計分析法。 我們不再僅僅滿足於 logit 的基本統計量，而是將其視為一個完整的統計分佈和時間序列來進行建模。我們系統性地引入了五類能夠從不同維度描述該序列特性的高階統計指標，包括：

分佈形狀特徵：透過計算偏度 (skewness) 與峰度 (kurtosis) 來捕捉 logit 分佈的不對稱性與集中趨勢。

資訊理論特徵：利用資訊熵 (entropy) 來量化模型在預測時的不確定性。

時序穩定性特徵：計算自相關係數 (autocorrelation) 來衡量 logit 序列隨時間變化的平滑程度。

分佈擬合特徵：採用高斯混合模型 (GMM) 來建模 logit 序列可能存在的多模態特性。

峰值穩健性特徵：透過極值理論中的 top-k 平均值來取代單一最大值，以獲得更可靠的峰值信心度。

我們將在公開的 SpeechOcean762 資料集 上驗證這些新指標的有效性，並與Parikh et al. 的基線方法進行深入比較。本研究期望能證明，透過對 logit 序列進行更深層次的統計建模，我們能夠更精準地捕捉到發音的細微差異，從而為自動化發音評估技術開闢新的可能性。

Electronically-available Resources

ROCLING provides this description and accompanying style files at ROCLING official website. We strongly recommend the use of these style files, which have been appropriately tailored for the ROCLING 2025 proceedings.

**MSWord-specific details**: All formatting is made available in the MSWord Styles in this template (rocling2025-final.docx). In newer versions of MSWord, click Home, then expand the Styles tile by clicking the diagonal arrow on the lower left corner. This should open all styles in the template for you to apply to your document as needed.

研究方法

本研究旨在透過對 logit 序列進行更深層次的統計分析，來提升發音錯誤檢測的準確性。在本章節中，我們首先將簡要回顧作為我們比較基準的 logit-based GOP 指標。接著，我們將詳細闡述本研究提出的五類多面向統計特徵，這些特徵旨在從分佈形狀、資訊量、時間穩定性等多個維度，更全面地捕捉發音的細微動態。

3.1. 基線 Logit-based GOP 指標 (Baseline Logit-based GOP Metrics)我們選用 Parikh et al. (2025) 所提出的主要 logit-based 指標作為效能比較的基線。這些指標代表了當前 logit-based GOP 方法的基礎。

最大 Logit (GOPMaxLogit​)：取音素對齊幀範圍內，目標音素 p 的 logit 序列 lt(p)​ 中的最大值。此指標反映了模型在整個發音過程中所達到的最高信心水準。

GOPMaxLogit​(p)=t∈[t1​,t2​]max​lt(p)​

Logit 邊界 (GOPMargin​)：計算在每一幀中，目標音素的 logit 值與最強競爭音素的 logit 值之間的差值，再將這些差值於整個音素段內取平均。此指標量化了目標音素在 logit 空間中的「突出程度」或「可區分性」。

GOPMargin​(p)=T1​t=t1​∑t2​​(lt(p)​−k=pmax​lt(k)​)

Logit 變異數 (GOPVarLogit​)：計算目標音素 logit 序列的變異數，用以衡量模型信心的穩定性。較低的變異數通常表示一個穩定、流暢的發音。

2.2. 提出的多面向統計指標 (Proposed Multi-faceted Statistical Metrics)

為了超越基線指標的侷限，我們引入了五類更為複雜的統計方法。這些方法被設計用來從 logit 序列中提取更深層次的資訊。

2.2.1. 分佈形狀特徵：動差分析 (Distribution Shape: Moment Analysis)

除了二階動差（變異數），更高階的動差能提供關於 logit 序列統計分佈「形狀」的額外資訊，這對於描述模型信心的動態變化至關重要。

偏度 (Skewness, G1​)：作為第三階標準化動差，偏度衡量 logit 分佈的不對稱性。正偏度可能表示模型信心是逐漸建立然後迅速下降的過程，而負偏度則相反。異常的偏斜可能暗示著不自然的發音模式。

峰度 (Kurtosis, G2​)：作為第四階標準化動差，峰度衡量分佈的「峰銳度」與「尾部厚度」。高峰度表示模型的信心高度集中於某個值，伴隨可能的極端離群值；低峰度則表示分佈較為平坦。這有助於識別發音過程中信心的集中或分散程度。

2.2.2. 資訊理論特徵：不確定性量化 (Information Theory: Uncertainty Quantification)

此方法從整個後驗機率分佈的角度出發，而非僅僅關注目標音素，用以量化模型在預測時的整體「混淆程度」。

平均夏農熵 (Mean Shannon Entropy)：我們計算音素段內每一幀的後驗機率分佈 P(xt​) 的夏農熵，然後取其平均值。熵是模型不確定性的直接度量。一個高的平均熵意味著模型的機率被分散在多個候選音素上，是發音含糊或錯誤的強烈信號。

Hmean​=T1​t=t1​∑t2​​(−k=1∑D​P(k∣xt​)logP(k∣xt​))

平均 KL 散度 (Mean KL Divergence)：此指標衡量每一幀的實際後驗機率分佈 P(xt​) 與一個代表「完美發音」的理想分佈（即目標音素機率為1的 one-hot 向量 Q）之間的「距離」。較大的 KL 散度意味著模型的輸出與理想狀態相去甚遠。

2.2.3. 分佈擬合特徵：高斯混合模型 (Distribution Fitting: Gaussian Mixture Models)

我們假設 logit 序列的分佈並非單峰，而是可能由多個潛在狀態（如音素的起始、穩定、結束階段）混合而成。高斯混合模型 (GMM) 能有效捕捉這種多模態特性。我們將 logit 序列擬合成一個包含 K 個高斯分量的 GMM，並提取其參數作為特徵，例如各分量的均值 (μk​)、變異數 (σk2​) 和權重 (wk​)。這些參數能精細地描述發音過程中模型信心的多階段動態。

2.2.4. 時序穩定性特徵：自相關分析 (Temporal Stability: Autocorrelation Analysis)

為了彌補現有方法忽略 logit 序列時間順序性的不足，我們引入時間序列分析。我們計算 logit 序列在延遲為 1 (lag-1) 時的自相關係數 (Autocorrelation)。一個高的正相關係數表示 logit 序列是平滑且穩定變化的，這通常對應於一個清晰、穩定的發音。反之，一個接近於零或負值的係數則暗示著序列存在劇烈、不規則的波動，可能是發音不穩定的跡象。

2.2.5. 峰值穩健性特徵：極值理論 (Peak Robustness: Extreme Value Theory)

$GOP\_{MaxLogit}$ 對單一的雜訊尖峰非常敏感。為了解決這個問題，我們採用一個更穩健的峰值估計方法：top-k 平均值。此方法選取 logit 序列中最大的 k 個值（例如 k=3），並計算它們的平均值。這提供了一個更穩定的模型「峰值信心」的估計，有效地平滑了單一離群值的影響。

3 實驗與結果 (Experiments & Results)

**3.1. 實驗設定 (Experimental Setup)**

**資料集 (Dataset):** 本研究的所有實驗皆於

**SpeechOcean762** 資料集上進行 。此資料集是一個專為發音評估研究所設計的公開語料庫，包含了由 250 位母語為中文的 L2 英語學習者所錄製的 5,000 句英語語音 。至關重要的是，該資料集提供了每個音素的標準音標 (canonical transcription) 和實際發音音標 (realized transcription) 。這讓我們可以透過程式化比對，客觀地產生「正確」或「錯誤」的標籤，作為我們分類任務的參考標準 (ground truth)，而無需依賴任何人類的主觀評分。

**評估指標 (Evaluation Metrics):** 我們將發音錯誤檢測定義為一個二元分類任務（正確 vs. 錯誤）。由於資料集中正確發音的樣本遠多於錯誤發音，存在類別不平衡問題，因此我們採用了以下指標來評估模型效能：

**準確率 (Accuracy)**、**精確率 (Precision)**、**召回率 (Recall)**、**F1-score** 以及 **馬修斯相關係數 (Matthews Correlation Coefficient, MCC)** 。其中，MCC 是一個在不平衡資料下表現特別穩健的分類效能指標，我們將其視為主要的評斷依據 。

**比較方法 (Compared Methods):** 我們將本研究提出的五類多面向統計指標，與 Parikh et al. (2025) 提出的

logit-based 基線指標進行比較，主要包括 $GOP\_{MaxLogit}$, $GOP\_{Margin}$ 和 $GOP\_{DNN}$ 。

**3.2. 結果與分析 (Results and Analysis)**

**3.2.1. 主要結果**

本實驗旨在評估多種 logit-based GOP 指標在發音錯誤檢測二元分類任務上的效能。詳細的實驗數據如表1所示，所有指標均以馬修斯相關係數 (MCC) 作為主要排序依據，因為它在處理類別不平衡的數據時最為穩健。

實驗結果揭示了幾個出乎意料的發現：

1. **整體分類效能有限**：首先，從數據中可以看出，所有單獨使用的 GOP 指標，在本次實驗中的分類能力都相對有限。表現最好的指標 kurtosis（峰度），其 MCC 分數也僅為 0.0819，這表明僅依靠單一維度的統計特徵，並設定一個最佳門檻值，難以完美地區分正確與錯誤的發音。
2. **高階動差表現最佳**：令人驚訝的是，在所有指標中，表現最好的是描述 logit 分佈**形狀**的高階動差特徵，特別是**峰度 (kurtosis)** 和**偏度 (skewness)**，它們在 MCC 排序中位列第一和第二。
3. **基線方法與部分新方法的權衡問題**：許多基線方法（如 mean\_logit\_margin）和我們提出的部分方法（如 evt\_k3, entropy\_mean）展現出極端的「高召回率、低精確率」現象。例如，gmm\_means\_0 和 entropy\_mean 的召回率達到了 1.0，但其準確率卻極低（約 0.12），MCC 分數也趨近於 0。這意味著，為了找到盡可能多的錯誤發音（高召回率），這些指標的最佳門檻值將幾乎所有的音素都標記為「錯誤」，從而導致了極高的誤報率（低精確率），使其作為分類器幾乎失效。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 發音錯誤檢測之分類效能比較 (以 MCC 分數排序) | | | | | |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

表格1：不同 GOP 指標於發音錯誤檢測之分類效能比較

venues, but will not be provided the list of previous presentations from the submission form.

|  |  |  |
| --- | --- | --- |
| **Type of Text** | **Font Size** | **Style** |
| paper title | 15 pt | **bold** |
| author names | 12 pt | **bold** |
| author affiliation | 12 pt |  |
| the word “Abstract” | 12 pt | **bold** |
| section titles | 12 pt | **bold** |
| document text | 11 pt |  |
| captions | 10 pt |  |
| abstract text | 10 pt |  |
| keyword text | 10 pt |  |
| bibliography | 10 pt |  |
| footnotes | 9 pt |  |

Table 1: Font guide.

Once a paper has been accepted to the conference, the camera-ready version of the paper should include the author's names and affiliations, and is allowed to use self-references.

**MS Word-specific details**: If the paper is accepted, remove the header, footer (page numbers), and the ruler for the final version (camera-ready).

Formatting Instructions

Manuscripts must be in two-column format. Exceptions to the two-column format include the title, authors' names and complete addresses, which must be centered at the top of the first page, and any full-width figures or tables (see the guidelines in Section 5.10). **Type single-spaced.** Start all pages directly under the top margin. The manuscript should be printed single-sided and its length should not exceed the maximum page limit described in Section 3. Pages should be numbered in the version submitted for review, but **pages should not be numbered in the camera-ready version**.

File Format

For the production of the electronic manuscript you must use Adobe's Portable Document Format (PDF). Please make sure that your PDF file includes all the necessary fonts (especially tree diagrams, symbols, and fonts with Asian characters). When you print or create the PDF file, there is usually an option in your printer setup to include none, all or just non-standard fonts. Please make sure that you select the option of including ALL the fonts. **Before sending it, test your PDF by printing it from a computer different from the one where it was created.** Moreover, some word processors may generate very large PDF files, where each page is rendered as an image. Such images may reproduce poorly. In this case, try alternative ways to obtain the PDF. One way on some systems is to install a driver for a postscript printer, send your document to the printer specifying ``Output to a file'', then convert the file to PDF.

It is of utmost importance to specify the **A4 format** (21 cm x 29.7 cm) when formatting the paper. Print-outs of the PDF file on A4 paper should be identical to the hardcopy version. If you cannot meet the above requirements about the production of your electronic submission, please contact the publication chairs as soon as possible.

**MS Word-specific details**: Lines should be justified, with even spacing between margins (Ctrl+J). Authors are encouraged to use Paragraph spacing at Multiple, 1.05pt, with Font character spacing condensed with kerning of 0.1pt, and Margins at 0.98in, for consistency with A4 paper and LaTeX-formatted documents. Go to Format, Document, Page Setup, and ensure A4 is selected.

* 1. Layout

Format manuscripts two columns to a page, in the manner these instructions are formatted. The exact dimensions for a page on A4 paper are:

* Left and right margins: 2.5 cm
* Top margin: 2.5 cm
* Bottom margin: 2.5 cm
* Column width: 7.7 cm
* Column height: 24.7 cm
* Gap between columns: 0.6 cm

Papers should not be submitted on any other paper size. If you cannot meet the above requirements about the production of your electronic submission, please contact the publication chairs above as soon as possible.

* 1. Fonts

For reasons of uniformity, Adobe's **Times Roman** font should be used. If Times Roman is not available, you may use the Times New Roman or Computer Modern Roman. In the Chinese manuscript, 標楷題 font should be used.

Table 1 specifies what font sizes and styles must be used for each type of text in the manuscript.

* 1. Ruler

A printed ruler (line numbers in the left and right margins of the article) should be presented in the version submitted for review, so that reviewers may comment on particular lines in the paper without circumlocution. The presence or absence of the ruler should not change the appearance of any other content on the page. The camera ready copy should not contain a ruler.

**Reviewers:** note that the ruler measurements may not align well with lines in the paper -- this turns out to be very difficult to do well when the paper contains many figures and equations, and, when done, looks ugly. In most cases one would expect that the approximate location will be adequate, although you can also use fractional references (e.g., this line ends at mark 291.5).

**MS Word-specific details**: In this Word template, you can add the ruler to each page by copying it from another page. The Ruler consists of separate left-hand rulers and right-hand rulers. Each ruler is a text box, sized 10 in x 0.45 in, flushed with the sides of the page, with a table containing numbers. The table has no borders, and the edges are pushed to the edges of the textbox. The left part of the ruler is justified left, left indent 0.1 in/right indent 0 in, and saved in the “ACL Ruler Left” style; and the right-hand ruler, justified right had left indent 0 in/right indent 0.1 in, saved in the “ACL Ruler Right” style. Number text is in Arial, and spacing between each line is Multiple, 1.05 in, with 4.5 pt below each line. After pasting a new ruler onto a page, highlight the ruler text and press **F9** to update the numbers. **Align** the text box **Middle**, and **Center**, and then **Arrange** Behind Text.

* 1. Title and Authors

Center the title, author's name(s) and affiliation(s) across both columns. Do not use footnotes for affiliations. Place the title centered at the top of the first page, in a 15-point bold font. Long titles should be typed on two lines without a blank line intervening. Put the title 2.5 cm from the top of the page, followed by a blank line, then the author's names(s), and the affiliation on the following line.

Do not use only initials for given names (middle initials are allowed). Do not format surnames in all capitals (e.g., use “Mitchell” not “MITCHELL”). Do not format title and section headings in all capitals except for proper names (such as “BLEU”) that are conventionally in all capitals. The affiliation should contain the author's complete address, and if possible, an electronic mail address.

The title, author names and addresses should be completely identical to those entered to the electronical paper submission website in order to maintain the consistency of author information among all publications of the conference. If they are different, the publication chairs may resolve the difference without consulting with you; so it is in your own interest to double-check that the information is consistent.

Start the body of the first page 7.5 cm from the top of the page. **Even in the anonymous version of the paper, you should maintain space for names and addresses so that they will fit in the final (accepted) version.**

* 1. Abstract

Use two-column format when you begin the abstract. Type the abstract at the beginning of the first column. The width of the abstract text should be smaller than the width of the columns for the text in the body of the paper by 0.6 cm on each side. Center the word **Abstract** in a 12 point bold font above the body of the abstract. The abstract should be a concise summary of the general thesis and conclusions of the paper. It should be no longer than 200 words. The abstract text should be in 10 point font.

In addition, for the Chinese manuscript, center the word 摘要 with the same format with English word Abstract, and the format of abstract is also the same to English abstract.

* 1. Text

Begin typing the main body of the text immediately after the abstract, observing the two-column format as shown in the present document.

Indent 0.4 cm when starting a new paragraph.

* 1. Sections

Format section and subsection headings in the style shown on the present document. Use numbered sections (Arabic numerals) to facilitate cross references. Number subsections with the section number and the subsection number separated by a dot, in Arabic numerals.

Footnotes

Put footnotes at the bottom of the page and use 9 point font. They may be numbered or referred to by asterisks or other symbols.[[1]](#footnote-1) Footnotes should be separated from the text by a line.[[2]](#footnote-2)

Graphics

Place figures, tables, and photographs in the paper near where they are first discussed, rather than at the end, if possible. Wide illustrations may run across both columns. Color is allowed, but adhere to Section 6's guidelines on accessibility.

**Captions:** Provide a caption for every illustration; number each one sequentially in the form: English manuscript uses

* “Figure 1. Caption of the Figure.”
* “Table 1. Caption of the Table.” .

And Chinese manuscript uses

* “圖1. Caption of the Figure.”
* “表1. Caption of the Table.” ,

Type the captions of the figures and tables below the body, using 10 point text. Captions should be placed below illustrations. Captions that are one line are centered (see Table 1). Captions longer than one line are left-aligned (see Figure 1).

**MS Word-specific details**:

**Creating:** To create a new Figure or Table, insert a Text Box where you want it to appear (generally, centered at the top of a column close to where it is referred to) and then fill it in with the Figure (or Table). Highlight and right click to add Caption, with the ACL Caption style (or ACL Caption Long style for multi-line captions), which places 10 pt below and above the caption.

**Numbering:** To update numbering, highlight all the relevant text (*e.g.*, **Ctrl-A + F9**). This will update all the numbering applicable to tables, figures, equations, and headings.

**Cross-referencing:** To add a cross reference to a figure or table:

* Place the mouse pointer at the location where you wish to add the cross-reference.
* Click on the **Insert** menu, (then click **Reference**), and then **Cross-reference** in the **Links** panel.
* In the **Cross-reference** dialog box, click the caption to which you are building the text reference.
* For a figure, under **Reference Type**, click **Figure**.
* Under Insert Reference To, click Only Label and Number, then click OK.

|  |
| --- |
|  |
| Figure 1: A figure with a caption that runs for more than one line**.** |

This is an example reference to Figure 1.

Hyperlinks

Within-document and external hyperlinks are indicated with Dark Blue text, Color Hex #000099.

Citations

Citations within the text appear in parentheses as ([Gusfield, 1997](#Gusfield1997)) or, if the author's name appears in the text itself, as Gusfield ([1997](#Gusfield1997)). Append lowercase letters to the year in cases of ambiguities. Treat double authors as in ([Aho and Ullman, 1972](#AhoUllman72)), but write as in ([Chandra et al., 1981](#ChandraEtAl1981)) when more than two authors are involved. Collapse multiple citations as in ([Gusfield, 1997](#Gusfield1997); [Aho and Ullman, 1972](#AhoUllman72)).

Refrain from using full citations as sentence constituents. Instead of

“([Gusfield, 1997](#Gusfield1997)) showed that ...”

or

“(Tseng et al., 2002) showed that ...”

, write

“Gusfield ([1997](#Gusfield1997)) showed that ...”

or

“Tseng et al. ([2002](#Tseng2002)) showed that ...”

References

Gather the full set of references together under the heading **References**; place the section before any Appendices. Arrange the references alphabetically by first author, rather than by order of occurrence in the text.

Provide as complete a citation as possible, using a consistent format, such as the one for *Computational Linguistics* or the one in the *Publication Manual of the American Psychological Association* ([American Psychological Association, 1983](#APA83)). Use full names for authors, not just initials.

Submissions should accurately reference prior and related work, including code and data. If a piece of prior work appeared in multiple venues, the version that appeared in a refereed, archival venue should be referenced. If multiple versions of a piece of prior work exist, the one used by the authors should be referenced. Authors should not rely on automated citation indices to provide accurate references for prior and related work.

The following text cites various types of articles so that the references section of the present document will include them.

* Example article in a journal: ([Chandra et al., 1981](#ChandraEtAl1981))
* Example article in proceedings, with location: ([Goodman et al., 2016](#GoodmanEtAl2016))
* Example article in proceedings, without location: ([Andrew and Gao., 20](#Andrew2007)07)
* Example arxiv paper: ([Rasooli and Tetreault, 2016](#Rasooli2015))

**MS Word-specific details:** To create hyperlinks between citations and references, as you insert each full reference in the References section, highlight it and then select Insert, Bookmark. Link back to the reference from its citations in the text by highlight the citation, right clicking, and selecting Insert, Cross-Reference, then selecting the Bookmark you’ve saved. Highlight the citation again to give make it dark blue (included in this theme), if it is not automatically applied. If there are problems saving the hyperlinks when you convert the document to PDF, use an online converter such as <http://go4convert.com>.

Digital Object Identifiers

All camera-ready references are required to contain the appropriate DOIs (or as a second resort, the hyperlinked ACL Anthology Identifier) to all cited works. Appropriate records should be found for most materials in the current ACL Anthology at <http://aclanthology.info/>. As examples, we cite ([Goodman et al., 2016](#GoodmanEtAl2016)) to show you how papers with a DOI will appear in the bibliography. We cite ([Harper, 2014](#Harper2014)) to show how papers without a DOI but with an ACL Anthology Identifier will appear in the bibliography.

* 1. Equations

An example equation is shown below:

(1)

To add new equations, authors are encouraged to copy this existing equation line, and then replace with the new equation. The numbering and alignment of equation line elements is automatic. To update equation numbering, press **Ctrl-A + F9**. Note: this will only update the number to the right of the equation; to update numbering within the text you must create a cross-reference.

**Cross-referencing:** To create a cross-reference for an equation:

* Create a bookmark for it.
* Select the number to the right of the equation. Go to **Insert**, **Bookmark** (in the **Links** panel),andthen create a name for your equation. Press **Add** to create the bookmark.
* To refer back, place the mouse pointer at the location where you wish to add the cross reference.
* Go to **Insert, Cross-reference** (in the **Links** panel).In the dialogue box, select **Bookmark** and **Bookmark Text** from each dropdown list. Uncheck **Insert as Hyperlink**, then click **OK**.
* This will make it such that whenever a new equation is added, the references to the equation will be updated when **Ctrl-A + F9** is pressed.
* This an example cross-reference to Equation 1.

1. Accessibility

In an effort to accommodate people who are color-blind (as well as those printing to paper), grayscale readability is strongly encouraged. Color is not forbidden, but authors should ensure that tables and figures do not rely solely on color to convey critical distinctions. A simple criterion: All curves and points in your figures should be clearly distinguishable without color.

1. MS Word STREAM Tools

This Microsoft Word file was updated in 2016 with STREAM Tools, designed for creating well-formatted reports and papers with Microsoft Word (Mamishev, 2010; Mamishev, 2013).

Acknowledgments

The acknowledgements should go immediately before the references. Do not number the acknowledgments section. Do not include this section when submitting your paper for review.

References

Alfred. V. Aho and Jeffrey D. Ullman. 1972. *The Theory of Parsing, Translation and Compiling, volume 1*. Prentice-Hall, Englewood Cliffs, NJ.

American Psychological Association. 1983. *Publications Manual.* American Psychological Association, Washington, DC.

Galen Andrew and Jianfeng Gao. 2007. Scalable training of L1-regularized log-linear models. In *Proceedings of the 24th International Conference on Machine Learning*, pages 33–40.

Ashok K. Chandra, Dexter C. Kozen, and Larry J. Stockmeyer. 1981. [Alternation](http://dl.acm.org/citation.cfm?doid=322234.322243). *Journal of the Association for Computing Machinery*, 28(1):114-133. https://doi.org/10.1145/322234.32224.

Association for Computing Machinery. 1983. *Computing Reviews*, 24(11):503-512.

James Goodman, Andreas Vlachos, and Jason Naradowsky. 2016. [Noise reduction and targeted exploration in imitation learning for abstract meaning representation parsing](http://aclweb.org/anthology/P16-1001). In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, pages 1–11. <https://doi.org/10.18653/v1/P16-1001>.

Dan Gusfield. 1997. *Algorithms on Strings, Trees and Sequences*. Cambridge University Press, Cambridge, UK.

Mary Harper. 2014. [Learning from 26 languages: Pro- gram management and science in the babel program](http://aclweb.org/anthology/C14-1001). In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*. Dublin City University and Association for Computational Linguistics, page 1. <http://aclweb.org/anthology/C14-1001>.

Alexander V. Mamishev and Murray Sargent. 2013. *Creating Research and Scientific Documents Using Microsoft Word*. Microsoft Press, Redmond, WA.

Alexander V. Mamishev and Sean D. Williams. 2010. *Technical Writing for Teams: The STREAM Tools Handbook*. Wiley-IEEE Press, Hoboken, NJ.

Mohammad Sadegh Rasooli and Joel R. Tetreault. 2015. *Yara parser: A fast and accurate depen-dency parser*. *Computing Research Repository,* arXiv:1503.06733. Version 2.

Hui-Hsin Tseng, Chao-Lin Liu, Zhao-Ming Gao, and Keh-Jiann Chen. 2002. 以構詞律與相似法為本的中文動詞自動分類研究 (a hybrid approach for automatic classification of Chinese unknown verbs) [in Chinese]. In *International Journal of Computational Linguistics & Chinese Language Processing, Volume 7, Number 1, February 2002: Special Issue on HowNet and Its Applications*, pages 1–28.

1. This is how a footnote should appear. [↑](#footnote-ref-1)
2. 2 Note the line separating the footnotes from the text. [↑](#footnote-ref-2)