

Organize PDF Files into a Dedicated Folder

```
import os
import shutil

target_folder = "Scanned_PDF_sources"
os.makedirs(target_folder, exist_ok=True)

# Moving only PDF files that do NOT contain "specific_test" in the filename
for file in os.listdir():
    if file.lower().endswith(".pdf") and "specific_test" not in file.lower():
        shutil.move(file, os.path.join(target_folder, file))

print(f"PDF files (excluding ones with 'specific_test' in the name) moved to '{target_folder}'")

PDF files (excluding ones with 'specific_test' in the name) moved to 'Scanned_PDF_sources'
```

Extract Contents of a ZIP File to Current Directory

```
import zipfile
import os

# Path to the zip file
zip_path = "Transcripts.zip"

# Destination folder where files will be extracted
extract_to = os.getcwd()

# Unzip the folder
with zipfile.ZipFile(zip_path, 'r') as zip_ref:
    zip_ref.extractall(extract_to)

print(f"Extracted to: {extract_to}")

Extracted to: d:\test_1
```

Convert First Three Pages of Each PDF to Images

```
import os
import fitz

# Defining input and output directories
input_folder = "Scanned_PDF_sources"
output_base_folder = "overall_folder"

# Creating output base folder if it doesn't exist
```

```

os.makedirs(output_base_folder, exist_ok=True)

def pdf_to_images(pdf_path, output_folder):
    pdf_name = os.path.splitext(os.path.basename(pdf_path))[0]
    pdf_output_folder = os.path.join(output_folder, pdf_name)
    os.makedirs(pdf_output_folder, exist_ok=True)

    doc = fitz.open(pdf_path)

    for i, page in enumerate(doc, start=1):
        if i <= 3: # Saving only the first 3 pages
            pix = page.get_pixmap()
            image_path = os.path.join(pdf_output_folder,
f"{pdf_name}_page_{i}.png")
            pix.save(image_path)

    print(f"Images saved in folder: {pdf_output_folder}")

# Processing all PDF files in the input folder
for filename in os.listdir(input_folder):
    if filename.lower().endswith(".pdf"):
        full_pdf_path = os.path.join(input_folder, filename)
        pdf_to_images(full_pdf_path, output_base_folder)

Images saved in folder: overall_folder\Buendia - Instruccion
Images saved in folder: overall_folder\Constituciones sinodales
Calahorra 1602
Images saved in folder: overall_folder\Ezcaray - Vozes
Images saved in folder: overall_folder\Mendo - Principe perfecto
Images saved in folder: overall_folder\Paredes - Reglas generales

```

Extract and Save Words and Coordinates from OCR with Margin-Based Filtering and Visualization

```

import cv2
import pytesseract
import pandas as pd
import matplotlib.pyplot as plt
from PIL import Image

# Input image path
image_path = "overall_folder\Buendia - Instruccion\Buendia - Instruccion_page_2.png" # Please adjust for testing for different documents

# Output directories
coordinates_dir = "Coordinates"
words_dir = "Words"
os.makedirs(coordinates_dir, exist_ok=True)
os.makedirs(words_dir, exist_ok=True)

```

```

# Extract base name
base_name = os.path.splitext(os.path.basename(image_path))[0]

# Load and prepare image
image = cv2.imread(image_path)
image_rgb = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)
pil_image = Image.fromarray(image_rgb)

# OCR with bounding box data
custom_config = r'--oem 3 --psm 3'
data = pytesseract.image_to_data(pil_image, config=custom_config,
output_type=pytesseract.Output.DICT)

# Converting to DataFrame
df = pd.DataFrame({
    "x1": data["left"],
    "y1": data["top"],
    "width": data["width"],
    "height": data["height"],
    "word": data["text"]
})

# Filtering out empty text
df = df[df["word"].str.strip() != ""]

# Computing bounding box (x2, y2)
df["x2"] = df["x1"] + df["width"]
df["y2"] = df["y1"] + df["height"]

# Margin metrics
df["left_margin_metric"] = df["y1"] + df["x1"]
df["right_margin_metric"] = df["x2"] + df["y2"]

bbox_largest_left_metric = df.loc[df["left_margin_metric"].idxmin()]
bbox_largest_right_metric = df.loc[df["right_margin_metric"].idxmax()]
left_margin = bbox_largest_left_metric["x1"]
right_margin = bbox_largest_right_metric["x2"]

# Filtering within margins
df["within_margins"] = (df["x2"] >= left_margin) & (df["x1"] <=
right_margin)
filtered_df = df[df["within_margins"]]

# Saving coordinates
coordinates_file = f"Coordinates/{base_name}_coord.txt"
with open(coordinates_file, "w") as f:
    for _, row in filtered_df.iterrows():
        f.write(f"{row['x1']},{row['y1']},{row['x2']},{row['y2']}\n")

```

```
# Saving words
words_file = f"Words/{base_name}_words.txt"
with open(words_file, "w") as f:
    for _, row in filtered_df.iterrows():
        f.write(f"{row['word']}\n")

# Drawing bounding boxes
for _, row in filtered_df.iterrows():
    cv2.rectangle(image_rgb, (row["x1"], row["y1"]), (row["x2"],
row["y2"]), (255, 0, 0), 2)

# Showing image with bounding boxes
plt.figure(figsize=(12, 12))
plt.imshow(image_rgb)
plt.title(f"Bounding Boxes: {base_name}")
plt.axis("off")
plt.show()

print(f"Processed image -> Coordinates: {coordinates_file}, Words:
{words_file}")
```

guro diseño de su edad: la Reli-
gion para con Dios en la devora
asistencia a los Templos; la piedad
con los Padres en la obediencia
mas rendida; y la modeltia, y de-
seo de saber, con los mayores,
gustando mas de oir, y pregun-
tar, que de definir, y resolver. Bien
que esto en vuestra infinita Sabi-
duria fué soberana dignacion, y
en la natural ignorancia de los
Niños es indispensable necesi-
dad.

Ni tienen solamente en Vos
el diseño, la luz, y el exemplo,
fino tambien el amor, y protec-
cion. Vos, como singular Maes-
tro de los Niños, les dais enten-
dimiento, y comunicais la sabi-
duria. Vos les prometeis el Reyno
de los Cielos, y os indignais con
quien les aparta de Vos, y les
proponéis por norma del can-
dor, inocencia, y christiana hu-
mildad. Vuestro amor parece que
no pudo explicarse mas tierno, y
liberal con los Niños, pues no
contento de echarles vuestras di-
divi-

divinas bendiciones, les unisteis
a vuestro sagrado pecho con in-
visibles abrazos. Dichosa edad,
que os mereció tan regalados ca-
rinos!

Y pues en la celestial Jeru-
salén no ha mudado de condicion
vuestra Benignidad, profeguid,
ó Niño tierno, y Dios Eterno,
profeguid en bendecirles, y favo-
recerles. Sean tan fervorolamen-
te devotos de vuestra Admirable
MADRE, que se porten como sus
hijos, y hermanos de leche con
Vos. Serán sabios, si fueren cas-
tos; que no entra vuestra Sabi-
duria, donde no ay mucha pure-
za de conciencia. Crezcan en
vuestro tanto temór, y amor, co-
mo en los años, y mucho mas.
Adelantense en la virtud, como
en las letras, y mucho mas, ha-
ca que lleguen, por vuestra imi-
tacion, a ser varones perfectos,
y consumados, agradables a
vuestros ojos, y provechosos a
la Republica, que libra casi co-
da su felicidad en la acertada
crian-

Luc. *ibid.*
Psal. 114. 6.
E 118. 130.
E 18. 8.
Matth. 19.
14.
Marc. 10.
14.
Matt. 18.
2. &c.

Marc. 10;
16.
Cant. 8. 1;
Sap. 1. 4;
Ad Ephes.
4. 13.

Processed image -> Coordinates: Coordinates/Buendia -
Instruccion_page_2_coord.txt, Words: Words/Buendia -
Instruccion_page_2_words.txt

Convert DOCX Transcripts to Word-Level TXT Files

```
import re
import os
from docx import Document

# Folder containing the .docx files
docx_folder = "Transcripts"

# Output folder for transcripts
transcripts_folder = "Transcript to txt"
os.makedirs(transcripts_folder, exist_ok=True)

# Process each .docx file in the folder
for file_name in os.listdir(docx_folder):
```

```

if file_name.endswith(".docx"): # Ensure it's a .docx file
    docx_path = os.path.join(docx_folder, file_name)
    doc = Document(docx_path)

    # Extract base name of the document (without extension)
    docx_name = os.path.splitext(file_name)[0]

    # Initialize variables
    page_content = {}
    current_page = None

    # Reading document and segregate by "PDF pX" headings
    for para in doc.paragraphs:
        text = para.text.strip()

        # Detecting headers like "PDF p1", "PDF p2", etc.
        if re.match(r"^PDF p\d+", text):
            current_page = text # Example: "PDF p1"
            page_content[current_page] = []
        elif current_page: # Store content under the detected
page header
            page_content[current_page].append(text)

    # Processing each page and save to corresponding .txt file
    for page, content in page_content.items():
        # Extract the page number from "PDF pX"
        page_number = re.search(r"\d+", page).group()

        # Generating output file name
        output_file = os.path.join(transcripts_folder,
f"{docx_name}_{page}_{page_number}_transcript.txt")

        # Flattening content into a single string
        full_text = " ".join(content)

        # Splitting into words (treating punctuation as separate
words)
        words = re.findall(r"[\w]+|[\^\w\s]", full_text)

        # Saving to file (one word per line)
        with open(output_file, "w", encoding="utf-8") as f:
            previous_word = None # To store the last word

            for word in words:
                if word == "-":
                    if previous_word is not None:
                        # Append '-' to the previous word instead
of adding a new line
                        previous_word += "-"
                    elif word == ",":

```

```

        if previous_word is not None:
            # Append '-' to the previous word instead
of adding a new line
            previous_word += ","

    else:
        if previous_word is not None:
            f.write(previous_word + "\n") # Write the
previous word

            previous_word = word # Update previous word

# Write the last word if it exists
if previous_word is not None:
    f.write(previous_word + "\n")

print(f"Saved: {output_file}")

print("Processing complete! Transcripts saved in the 'Transcripts'
folder.")

```

```

Saved: Transcript to txt\Buendia transcription_page_1_transcript.txt
Saved: Transcript to txt\Buendia transcription_page_2_transcript.txt
Saved: Transcript to txt\Buendia transcription_page_3_transcript.txt
Saved: Transcript to txt\Constituciones sinodales
transcription_page_1_transcript.txt
Saved: Transcript to txt\Constituciones sinodales
transcription_page_2_transcript.txt
Saved: Transcript to txt\Constituciones sinodales
transcription_page_3_transcript.txt
Saved: Transcript to txt\Ezcaray transcription_page_1_transcript.txt
Saved: Transcript to txt\Ezcaray transcription_page_2_transcript.txt
Saved: Transcript to txt\Ezcaray transcription_page_3_transcript.txt
Saved: Transcript to txt\Mendo transcription_page_1_transcript.txt
Saved: Transcript to txt\Mendo transcription_page_2_transcript.txt
Saved: Transcript to txt\Mendo transcription_page_3_transcript.txt
Saved: Transcript to txt\Mendo transcription_page_4_transcript.txt
Saved: Transcript to txt\Mendo transcription_page_5_transcript.txt
Saved: Transcript to txt\Paredes transcription_page_1_transcript.txt
Saved: Transcript to txt\Paredes transcription_page_2_transcript.txt
Saved: Transcript to txt\Paredes transcription_page_3_transcript.txt
Saved: Transcript to txt\PORCONES.228.35 1636
transcription_page_1_transcript.txt
Saved: Transcript to txt\PORCONES.228.35 1636
transcription_page_2_transcript.txt
Saved: Transcript to txt\PORCONES.228.35 1636
transcription_page_3_transcript.txt
Processing complete! Transcripts saved in the 'Transcripts' folder.

```


Crop Word Images Using Bounding Boxes and Correct Filenames via Levenshtein Matching with Transcript

```
import cv2
import os

# Load the image
image_path = "overall_folder\Buendia - Instruccion\Buendia - Instruccion_page_2.png" # Update according to the required document
image = cv2.imread(image_path)

# Extract base name for saving
image_name = os.path.splitext(os.path.basename(image_path))[0]

# Directory to save cropped word images
output_dir = f"{image_name}_cropped_words"
os.makedirs(output_dir, exist_ok=True)

# Read coordinates from the file
coordinates_file = f"Coordinates\{image_name}_coord.txt"
with open(coordinates_file, "r") as f:
    lines = f.readlines()

# Read words from the text file
words_file = f"Words\{image_name}_words.txt"
with open(words_file, "r", encoding="utf-8") as f:
    words = [line.strip() for line in f.readlines() if line.strip()]
# Remove empty lines

# Ensure we have the same number of words and coordinates
if len(words) != len(lines):
    print(f"Warning: Mismatch in the number of words ({len(words)}) and coordinates ({len(lines)}).")

# Iterate through coordinates and crop images
for idx, line in enumerate(lines):
    x1, y1, x2, y2 = map(int, line.strip().split(","))

    # Crop the word from the image
    cropped_word = image[y1:y2, x1:x2]

    # Get the corresponding word from the list
    word_name = words[idx] if idx < len(words) else f"word_{idx+1}"
    #sanitized_word_name = "".join(c for c in word_name if c.isalnum() or c in (' ', '_')) # Clean filename

    # Save the cropped image
    cropped_image_path = os.path.join(output_dir, f"{word_name}.png")
    cv2.imwrite(cropped_image_path, cropped_word)
    print(f"Saved cropped word '{word_name}' at {cropped_image_path}")
```



```

print("Cropping complete. All word images are saved in the
'cropped_words_pg_1' directory.")

#####

#####

# Performing corection of image_name

import os
import shutil
import Levenshtein

# Paths
transcript_file = "Transcript to txt\Buendia
transcription_page_2_transcript.txt"
input_folder = f"{image_name}_cropped_words"

input_folder_name = os.path.splitext(os.path.basename(input_folder))
[0]

output_folder = f"{input_folder_name}_corrected"

# Ensure the output directory exists
os.makedirs(output_folder, exist_ok=True)

# Function to load words from the transcript (preserving case)
def load_transcript_words(file_path):
    with open(file_path, "r", encoding="utf-8") as f:
        words = set(f.read().split()) # Extract unique words while
preserving case
    return words

# Function to find the most similar word while keeping capitalization
def find_best_match(word, word_list):
    best_match = None
    best_score = float("inf")

    for transcript_word in word_list:
        score = Levenshtein.distance(word.lower(),
transcript_word.lower()) # Compute Levenshtein distance
        if score < best_score:
            best_score = score
            best_match = transcript_word # Preserve original
capitalization

    return best_match if best_score <= len(word) * 0.4 else None #

```

```

Only accept if similarity is above 60%

# Load words from transcript (preserving case)
transcript_words = load_transcript_words(transcript_file)

# Keep track of used names to remove duplicates
used_filenames = set()

# Process each image in the input folder
for image_file in os.listdir(input_folder):
    input_image_path = os.path.join(input_folder, image_file)

    # Extract the base name (removing extension)
    image_name, ext = os.path.splitext(image_file)

    # Find the best match while keeping capitalization
    best_match = find_best_match(image_name, transcript_words)

    if best_match:
        new_filename = f"{best_match}.png" # Keep capitalization
        output_image_path = os.path.join(output_folder, new_filename)

        # If this word is already assigned, do not save a duplicate
        if best_match in used_filenames:
            print(f"Skipped duplicate: {image_file} (Already assigned to {new_filename})")
        else:
            # Copy the image to the new folder with the corrected name
            shutil.copy(input_image_path, output_image_path)
            used_filenames.add(best_match)
            print(f"Saved: {image_file} -> {output_image_path}")
    else:
        print(f"Skipped: {image_file} (No suitable match found)")

print("Processing complete! All corrected images are saved in 'cropped_words_pg_1_corrected'.")

Saved cropped word 'guro' at Buendia -
Instruccion_page_2_cropped_words\guro.png
Saved cropped word 'diffeiio' at Buendia -
Instruccion_page_2_cropped_words\diffeiio.png
Saved cropped word 'ag' at Buendia - Instruccion_page_2_cropped_words\
ag.png
Saved cropped word 'edad' at Buendia -
Instruccion_page_2_cropped_words\edad.png
Saved cropped word ':' at Buendia -
Instruccion_page_2_cropped_words\:.png
Saved cropped word 'Ia' at Buendia - Instruccion_page_2_cropped_words\
Ia.png
Saved cropped word 'Reli-' at Buendia -

```

Instruccion_page_2_cropped_words\Reli-.png
Saved cropped word 'gion' at Buendia -
Instruccion_page_2_cropped_words\gion.png
Saved cropped word 'para' at Buendia -
Instruccion_page_2_cropped_words\para.png
Saved cropped word 'con' at Buendia -
Instruccion_page_2_cropped_words\con.png
Saved cropped word 'Dios' at Buendia -
Instruccion_page_2_cropped_words\Dios.png
Saved cropped word 'en' at Buendia - Instruccion_page_2_cropped_words\
en.png
Saved cropped word 'la' at Buendia - Instruccion_page_2_cropped_words\
la.png
Saved cropped word 'devora' at Buendia -
Instruccion_page_2_cropped_words\devora.png
Saved cropped word 'afsiftécia' at Buendia -
Instruccion_page_2_cropped_words\afsiftécia.png
Saved cropped word '4' at Buendia - Instruccion_page_2_cropped_words\
4.png
Saved cropped word 'los' at Buendia -
Instruccion_page_2_cropped_words\los.png
Saved cropped word 'Templos;la' at Buendia -
Instruccion_page_2_cropped_words\Templos;la.png
Saved cropped word 'piedad' at Buendia -
Instruccion_page_2_cropped_words\piedad.png
Saved cropped word 'con' at Buendia -
Instruccion_page_2_cropped_words\con.png
Saved cropped word 'los' at Buendia -
Instruccion_page_2_cropped_words\los.png
Saved cropped word 'Padres' at Buendia -
Instruccion_page_2_cropped_words\Padres.png
Saved cropped word 'en' at Buendia - Instruccion_page_2_cropped_words\
en.png
Saved cropped word 'la' at Buendia - Instruccion_page_2_cropped_words\
la.png
Saved cropped word 'obediencia' at Buendia -
Instruccion_page_2_cropped_words\obediencia.png
Saved cropped word ''mas' at Buendia -
Instruccion_page_2_cropped_words\'mas.png
Saved cropped word 'rendida;' at Buendia -
Instruccion_page_2_cropped_words\rendida;.png
Saved cropped word '¥:' at ¥: .png
Saved cropped word 'modettia,' at Buendia -
Instruccion_page_2_cropped_words\modettia,.png
Saved cropped word 'y' at Buendia - Instruccion_page_2_cropped_words\
y.png
Saved cropped word 'de=' at Buendia -
Instruccion_page_2_cropped_words\de=.png
Saved cropped word 'feo' at Buendia -

Instruccion_page_2_cropped_words\feo.png
Saved cropped word 'de' at Buendia - Instruccion_page_2_cropped_words\
de.png
Saved cropped word 'faber' at Buendia -
Instruccion_page_2_cropped_words\faber.png
Saved cropped word 'con' at Buendia -
Instruccion_page_2_cropped_words\con.png
Saved cropped word 'los' at Buendia -
Instruccion_page_2_cropped_words\los.png
Saved cropped word 'mayores,' at Buendia -
Instruccion_page_2_cropped_words\mayores,.png
Saved cropped word 'guftando' at Buendia -
Instruccion_page_2_cropped_words\guftando.png
Saved cropped word 'mas' at Buendia -
Instruccion_page_2_cropped_words\mas.png
Saved cropped word 'de' at Buendia - Instruccion_page_2_cropped_words\
de.png
Saved cropped word 'oir,' at Buendia -
Instruccion_page_2_cropped_words\oir,.png
Saved cropped word 'y' at Buendia - Instruccion_page_2_cropped_words\
y.png
Saved cropped word 'pregun-' at Buendia -
Instruccion_page_2_cropped_words\pregun-.png
Saved cropped word 'tar,' at Buendia -
Instruccion_page_2_cropped_words\tar,.png
Saved cropped word 'que' at Buendia -
Instruccion_page_2_cropped_words\que.png
Saved cropped word 'de' at Buendia - Instruccion_page_2_cropped_words\
de.png
Saved cropped word 'definir,y' at Buendia -
Instruccion_page_2_cropped_words\definir,y.png
Saved cropped word 'refolver.Bien' at Buendia -
Instruccion_page_2_cropped_words\refolver.Bien.png
Saved cropped word 'que' at Buendia -
Instruccion_page_2_cropped_words\que.png
Saved cropped word 'efto' at Buendia -
Instruccion_page_2_cropped_words\efto.png
Saved cropped word 'en' at Buendia - Instruccion_page_2_cropped_words\
en.png
Saved cropped word 'vueftra' at Buendia -
Instruccion_page_2_cropped_words\vueftra.png
Saved cropped word 'infinita.' at Buendia -
Instruccion_page_2_cropped_words\infinita..png
Saved cropped word 'Sabi-,' at Buendia -
Instruccion_page_2_cropped_words\Sabi-,.png
Saved cropped word 'duria' at Buendia -
Instruccion_page_2_cropped_words\duria.png
Saved cropped word 'fue' at Buendia -
Instruccion_page_2_cropped_words\fue.png

Saved cropped word 'foberana' at Buendia -
Instruccion_page_2_cropped_words\foberana.png
Saved cropped word 'dignacion' at Buendia -
Instruccion_page_2_cropped_words\dignacion.png
Saved cropped word ',' at Buendia -
Instruccion_page_2_cropped_words\,.png
Saved cropped word 'y' at Buendia - Instruccion_page_2_cropped_words\
y.png
Saved cropped word 'en' at Buendia - Instruccion_page_2_cropped_words\
en.png
Saved cropped word 'la' at Buendia - Instruccion_page_2_cropped_words\
la.png
Saved cropped word 'natural' at Buendia -
Instruccion_page_2_cropped_words\natural.png
Saved cropped word 'ignorancia' at Buendia -
Instruccion_page_2_cropped_words\ignorancia.png
Saved cropped word 'de' at Buendia - Instruccion_page_2_cropped_words\
de.png
Saved cropped word 'los' at Buendia -
Instruccion_page_2_cropped_words\los.png
Saved cropped word 'Nifios' at Buendia -
Instruccion_page_2_cropped_words\Nifios.png
Saved cropped word 'es' at Buendia - Instruccion_page_2_cropped_words\
es.png
Saved cropped word 'indifpenfable' at Buendia -
Instruccion_page_2_cropped_words\indifpenfable.png
Saved cropped word 'necefsi-' at Buendia -
Instruccion_page_2_cropped_words\necefsi-.png
Saved cropped word 'dad.' at Buendia -
Instruccion_page_2_cropped_words\dad..png
Saved cropped word 'Ni' at Buendia - Instruccion_page_2_cropped_words\
Ni.png
Saved cropped word 'tienen' at Buendia -
Instruccion_page_2_cropped_words\tienen.png
Saved cropped word 'folamente' at Buendia -
Instruccion_page_2_cropped_words\folamente.png
Saved cropped word 'en' at Buendia - Instruccion_page_2_cropped_words\
en.png
Saved cropped word 'Vos' at Buendia -
Instruccion_page_2_cropped_words\Vos.png
Saved cropped word 'el' at Buendia - Instruccion_page_2_cropped_words\
el.png
Saved cropped word 'diffefio' at Buendia -
Instruccion_page_2_cropped_words\diffefio.png
Saved cropped word ',' at Buendia -
Instruccion_page_2_cropped_words\,.png
Saved cropped word 'laluz,y' at Buendia -
Instruccion_page_2_cropped_words\laluz,y.png
Saved cropped word 'el' at Buendia - Instruccion_page_2_cropped_words\
el.png

el.png
Saved cropped word 'exemplo,' at Buendia -
Instruccion_page_2_cropped_words\exemplo,.png
Saved cropped word 'fino' at Buendia -
Instruccion_page_2_cropped_words\fino.png
Saved cropped word 'tamibien' at Buendia -
Instruccion_page_2_cropped_words\tamibien.png
Saved cropped word 'el' at Buendia - Instruccion_page_2_cropped_words\
el.png
Saved cropped word 'amor,' at Buendia -
Instruccion_page_2_cropped_words\amor,.png
Saved cropped word 'y' at Buendia - Instruccion_page_2_cropped_words\
y.png
Saved cropped word 'protec-' at Buendia -
Instruccion_page_2_cropped_words\protec-.png
Saved cropped word 'cion.' at Buendia -
Instruccion_page_2_cropped_words\cion..png
Saved cropped word 'Vos,' at Buendia -
Instruccion_page_2_cropped_words\Vos,.png
Saved cropped word 'como' at Buendia -
Instruccion_page_2_cropped_words\como.png
Saved cropped word 'fingularsMaef=' at Buendia -
Instruccion_page_2_cropped_words\fingularsMaef=.png
Saved cropped word 'tro' at Buendia -
Instruccion_page_2_cropped_words\tro.png
Saved cropped word 'dé' at Buendia - Instruccion_page_2_cropped_words\
dé.png
Saved cropped word 'los' at Buendia -
Instruccion_page_2_cropped_words\los.png
Saved cropped word 'Nifios' at Buendia -
Instruccion_page_2_cropped_words\Nifios.png
Saved cropped word ',' at Buendia -
Instruccion_page_2_cropped_words\,.png
Saved cropped word 'les' at Buendia -
Instruccion_page_2_cropped_words\les.png
Saved cropped word 'dais' at Buendia -
Instruccion_page_2_cropped_words\dais.png
Saved cropped word 'enten-' at Buendia -
Instruccion_page_2_cropped_words\enten-.png
Saved cropped word 'dimiento' at Buendia -
Instruccion_page_2_cropped_words\dimiento.png
Saved cropped word ',' at Buendia -
Instruccion_page_2_cropped_words\,.png
Saved cropped word 'y' at Buendia - Instruccion_page_2_cropped_words\
y.png
Saved cropped word 'comunica's' at Buendia -
Instruccion_page_2_cropped_words\comunica's.png
Saved cropped word 'la' at Buendia - Instruccion_page_2_cropped_words\
la.png

Saved cropped word 'fabi-' at Buendia -
Instruccion_page_2_cropped_words\fabi-.png
Saved cropped word 'duria.Vos' at Buendia -
Instruccion_page_2_cropped_words\duria.Vos.png
Saved cropped word 'les' at Buendia -
Instruccion_page_2_cropped_words\les.png
Saved cropped word 'prometeis' at Buendia -
Instruccion_page_2_cropped_words\prometeis.png
Saved cropped word 'el' at Buendia - Instruccion_page_2_cropped_words\
el.png
Saved cropped word 'Reyno.' at Buendia -
Instruccion_page_2_cropped_words\Reyno..png
Saved cropped word ''de' at Buendia -
Instruccion_page_2_cropped_words\'de.png
Saved cropped word 'los' at Buendia -
Instruccion_page_2_cropped_words\los.png
Saved cropped word 'Cielos' at Buendia -
Instruccion_page_2_cropped_words\Cielos.png
Saved cropped word ', ' at Buendia -
Instruccion_page_2_cropped_words\,.png
Saved cropped word 'y' at Buendia - Instruccion_page_2_cropped_words\
y.png
Saved cropped word 'os' at Buendia - Instruccion_page_2_cropped_words\
os.png
Saved cropped word 'indignais' at Buendia -
Instruccion_page_2_cropped_words\indignais.png
Saved cropped word 'con' at Buendia -
Instruccion_page_2_cropped_words\con.png
Saved cropped word 'quien' at Buendia -
Instruccion_page_2_cropped_words\quien.png
Saved cropped word 'les' at Buendia -
Instruccion_page_2_cropped_words\les.png
Saved cropped word 'aparta' at Buendia -
Instruccion_page_2_cropped_words\aparta.png
Saved cropped word 'de' at Buendia - Instruccion_page_2_cropped_words\
de.png
Saved cropped word 'Vos,' at Buendia -
Instruccion_page_2_cropped_words\Vos,.png
Saved cropped word 'y' at Buendia - Instruccion_page_2_cropped_words\
y.png
Saved cropped word 'les' at Buendia -
Instruccion_page_2_cropped_words\les.png
Saved cropped word 'proponeis' at Buendia -
Instruccion_page_2_cropped_words\proponeis.png
Saved cropped word 'por' at Buendia -
Instruccion_page_2_cropped_words\por.png
Saved cropped word 'norma' at Buendia -
Instruccion_page_2_cropped_words\norma.png
Saved cropped word 'del'' at Buendia -

Instruccion_page_2_cropped_words\del'.png
Saved cropped word 'can-' at Buendia -
Instruccion_page_2_cropped_words\can-.png
Saved cropped word 'dor' at Buendia -
Instruccion_page_2_cropped_words\dor.png
Saved cropped word ', ' at Buendia -
Instruccion_page_2_cropped_words\,.png
Saved cropped word 'inocencia' at Buendia -
Instruccion_page_2_cropped_words\inocencia.png
Saved cropped word ', ' at Buendia -
Instruccion_page_2_cropped_words\,.png
Saved cropped word 'y' at Buendia - Instruccion_page_2_cropped_words\
y.png
Saved cropped word 'chription' at Buendia -
Instruccion_page_2_cropped_words\chription.png
Saved cropped word 'hu-' at Buendia -
Instruccion_page_2_cropped_words\hu-.png
Saved cropped word '2' at Buendia - Instruccion_page_2_cropped_words\
2.png
Saved cropped word '»y' at Buendia -
Instruccion_page_2_cropped_words\»y.png
Saved cropped word '«mildad.' at Buendia -
Instruccion_page_2_cropped_words\«mildad..png
Saved cropped word 'Vueftro' at Buendia -
Instruccion_page_2_cropped_words\Vueftro.png
Saved cropped word 'amor' at Buendia -
Instruccion_page_2_cropped_words\amor.png
Saved cropped word 'parece' at Buendia -
Instruccion_page_2_cropped_words\parece.png
Saved cropped word 'que' at Buendia -
Instruccion_page_2_cropped_words\que.png
Saved cropped word 'no' at Buendia - Instruccion_page_2_cropped_words\
no.png
Saved cropped word 'pudo' at Buendia -
Instruccion_page_2_cropped_words\pudo.png
Saved cropped word 'explicarfe' at Buendia -
Instruccion_page_2_cropped_words\explicarfe.png
Saved cropped word 'mas' at Buendia -
Instruccion_page_2_cropped_words\mas.png
Saved cropped word 'tierno,'y' at Buendia -
Instruccion_page_2_cropped_words\tierno,'y.png
Saved cropped word 'liberal' at Buendia -
Instruccion_page_2_cropped_words\liberal.png
Saved cropped word 'con' at Buendia -
Instruccion_page_2_cropped_words\con.png
Saved cropped word 'los' at Buendia -
Instruccion_page_2_cropped_words\los.png
Saved cropped word 'Nifios' at Buendia -
Instruccion_page_2_cropped_words\Nifios.png

Saved cropped word ',' at Buendia -
Instruccion_page_2_cropped_words\,.png
Saved cropped word 'pues' at Buendia -
Instruccion_page_2_cropped_words\pues.png
Saved cropped word 'no' at Buendia - Instruccion_page_2_cropped_words\
no.png
Saved cropped word 'contento' at Buendia -
Instruccion_page_2_cropped_words\contento.png
Saved cropped word 'de' at Buendia - Instruccion_page_2_cropped_words\
de.png
Saved cropped word 'echarles' at Buendia -
Instruccion_page_2_cropped_words\echarles.png
Saved cropped word 'vueftras' at Buendia -
Instruccion_page_2_cropped_words\vueftras.png
Saved cropped word 'dis' at Buendia -
Instruccion_page_2_cropped_words\dis.png
Saved cropped word 'divinas' at Buendia -
Instruccion_page_2_cropped_words\divinas.png
Saved cropped word 'bendiciones.,' at Buendia -
Instruccion_page_2_cropped_words\bendiciones.,.png
Saved cropped word 'les' at Buendia -
Instruccion_page_2_cropped_words\les.png
Saved cropped word 'unifteis' at Buendia -
Instruccion_page_2_cropped_words\unifteis.png
Saved cropped word 'a' at Buendia - Instruccion_page_2_cropped_words\
a.png
Saved cropped word 'vueftro' at Buendia -
Instruccion_page_2_cropped_words\vueftro.png
Saved cropped word 'fagrado' at Buendia -
Instruccion_page_2_cropped_words\fagrado.png
Saved cropped word 'pecho' at Buendia -
Instruccion_page_2_cropped_words\pecho.png
Saved cropped word 'con' at Buendia -
Instruccion_page_2_cropped_words\con.png
Saved cropped word 'f{ua-' at Buendia -
Instruccion_page_2_cropped_words\f{ua-.png
Saved cropped word 'vifsimos' at Buendia -
Instruccion_page_2_cropped_words\vifsimos.png
Saved cropped word 'abrazos.' at Buendia -
Instruccion_page_2_cropped_words\abrazos..png
Saved cropped word 'Dichofa' at Buendia -
Instruccion_page_2_cropped_words\Dichofa.png
Saved cropped word 'edad,' at Buendia -
Instruccion_page_2_cropped_words\edad,.png
Saved cropped word 'que' at Buendia -
Instruccion_page_2_cropped_words\que.png
Saved cropped word 'os' at Buendia - Instruccion_page_2_cropped_words\
os.png
Saved cropped word 'merecid' at Buendia -

Instruccion_page_2_cropped_words\merecid.png
Saved cropped word 'tan' at Buendia -
Instruccion_page_2_cropped_words\tan.png
Saved cropped word 'regalados' at Buendia -
Instruccion_page_2_cropped_words\regalados.png
Saved cropped word 'ca-' at Buendia -
Instruccion_page_2_cropped_words\ca-.png
Saved cropped word 'rifios' at Buendia -
Instruccion_page_2_cropped_words\rifios.png
Saved cropped word '!' at Buendia -
Instruccion_page_2_cropped_words\!.png
Saved cropped word '-' at Buendia -
Instruccion_page_2_cropped_words\-.png
Saved cropped word '~' at Buendia - Instruccion_page_2_cropped_words\
~.png
Saved cropped word 'Y' at Buendia - Instruccion_page_2_cropped_words\
Y.png
Saved cropped word 'pues' at Buendia -
Instruccion_page_2_cropped_words\pues.png
Saved cropped word 'en' at Buendia - Instruccion_page_2_cropped_words\
en.png
Saved cropped word 'la' at Buendia - Instruccion_page_2_cropped_words\
la.png
Saved cropped word 'celeftial' at Buendia -
Instruccion_page_2_cropped_words\celeftial.png
Saved cropped word 'Jeru-' at Buendia -
Instruccion_page_2_cropped_words\Jeru-.png
Saved cropped word 'falén' at Buendia -
Instruccion_page_2_cropped_words\falén.png
Saved cropped word 'no' at Buendia - Instruccion_page_2_cropped_words\
no.png
Saved cropped word 'ha' at Buendia - Instruccion_page_2_cropped_words\
ha.png
Saved cropped word 'mudado' at Buendia -
Instruccion_page_2_cropped_words\mudado.png
Saved cropped word 'de' at Buendia - Instruccion_page_2_cropped_words\
de.png
Saved cropped word 'condicion' at Buendia -
Instruccion_page_2_cropped_words\condicion.png
Saved cropped word 'vuettra' at Buendia -
Instruccion_page_2_cropped_words\vuettra.png
Saved cropped word 'Benignidad' at Buendia -
Instruccion_page_2_cropped_words\Benignidad.png
Saved cropped word ',' at Buendia -
Instruccion_page_2_cropped_words\,.png
Saved cropped word 'profeguid,' at Buendia -
Instruccion_page_2_cropped_words\profeguid,.png
Saved cropped word '6' at Buendia - Instruccion_page_2_cropped_words\
6.png

Saved cropped word 'Nifio' at Buendia -
Instruccion_page_2_cropped_words\Nifio.png
Saved cropped word 'tierno,' at Buendia -
Instruccion_page_2_cropped_words\tierno,.png
Saved cropped word 'y' at Buendia - Instruccion_page_2_cropped_words\
y.png
Saved cropped word 'Dios' at Buendia -
Instruccion_page_2_cropped_words\Dios.png
Saved cropped word 'Eterno,' at Buendia -
Instruccion_page_2_cropped_words\Eterno,.png
Saved cropped word 'profeguid' at Buendia -
Instruccion_page_2_cropped_words\profeguid.png
Saved cropped word 'en' at Buendia - Instruccion_page_2_cropped_words\
en.png
Saved cropped word 'bendecirles,' at Buendia -
Instruccion_page_2_cropped_words\bendecirles,.png
Saved cropped word 'y' at Buendia - Instruccion_page_2_cropped_words\
y.png
Saved cropped word 'favo-' at Buendia -
Instruccion_page_2_cropped_words\favo-.png
Saved cropped word 'recerles.' at Buendia -
Instruccion_page_2_cropped_words\recerles..png
Saved cropped word 'Sean' at Buendia -
Instruccion_page_2_cropped_words\Sean.png
Saved cropped word 'tan' at Buendia -
Instruccion_page_2_cropped_words\tan.png
Saved cropped word 'fervorofamen-' at Buendia -
Instruccion_page_2_cropped_words\fervorofamen-.png
Saved cropped word 'te' at Buendia - Instruccion_page_2_cropped_words\
te.png
Saved cropped word 'devotos' at Buendia -
Instruccion_page_2_cropped_words\devotos.png
Saved cropped word 'de' at Buendia - Instruccion_page_2_cropped_words\
de.png
Saved cropped word 'vueftra' at Buendia -
Instruccion_page_2_cropped_words\vueftra.png
Saved cropped word 'Admirable' at Buendia -
Instruccion_page_2_cropped_words\Admirable.png
Saved cropped word 'Manpre,' at Buendia -
Instruccion_page_2_cropped_words\Manpre,.png
Saved cropped word 'que' at Buendia -
Instruccion_page_2_cropped_words\que.png
Saved cropped word 'fe' at Buendia - Instruccion_page_2_cropped_words\
fe.png
Saved cropped word 'porten' at Buendia -
Instruccion_page_2_cropped_words\porten.png
Saved cropped word 'como' at Buendia -
Instruccion_page_2_cropped_words\como.png
Saved cropped word 'fus' at Buendia -

Instruccion_page_2_cropped_words\fus.png
Saved cropped word 'hijos' at Buendia -
Instruccion_page_2_cropped_words\hijos.png
Saved cropped word ',' at Buendia -
Instruccion_page_2_cropped_words\,.png
Saved cropped word 'y' at Buendia - Instruccion_page_2_cropped_words\
y.png
Saved cropped word 'hermanos' at Buendia -
Instruccion_page_2_cropped_words\hermanos.png
Saved cropped word 'de' at Buendia - Instruccion_page_2_cropped_words\
de.png
Saved cropped word 'leche' at Buendia -
Instruccion_page_2_cropped_words\leche.png
Saved cropped word 'con' at Buendia -
Instruccion_page_2_cropped_words\con.png
Saved cropped word 'Vos.' at Buendia -
Instruccion_page_2_cropped_words\Vos..png
Saved cropped word 'Seran' at Buendia -
Instruccion_page_2_cropped_words\Seran.png
Saved cropped word 'fabios,' at Buendia -
Instruccion_page_2_cropped_words\fabios,.png
Saved cropped word 'fi' at Buendia - Instruccion_page_2_cropped_words\
fi.png
Saved cropped word 'fueren' at Buendia -
Instruccion_page_2_cropped_words\fueren.png
Saved cropped word 'caf-' at Buendia -
Instruccion_page_2_cropped_words\caf-.png
Saved cropped word 'tos' at Buendia -
Instruccion_page_2_cropped_words\tos.png
Saved cropped word ';' at Buendia -
Instruccion_page_2_cropped_words\;.png
Saved cropped word 'que' at Buendia -
Instruccion_page_2_cropped_words\que.png
Saved cropped word 'no' at Buendia - Instruccion_page_2_cropped_words\
no.png
Saved cropped word 'entra' at Buendia -
Instruccion_page_2_cropped_words\entra.png
Saved cropped word 'vueftra' at Buendia -
Instruccion_page_2_cropped_words\vueftra.png
Saved cropped word 'Sabi-' at Buendia -
Instruccion_page_2_cropped_words\Sabi-.png
Saved cropped word ''' at Buendia -
Instruccion_page_2_cropped_words\'.png
Saved cropped word 'duria' at Buendia -
Instruccion_page_2_cropped_words\duria.png
Saved cropped word ',donde' at Buendia -
Instruccion_page_2_cropped_words\,donde.png
Saved cropped word 'no' at Buendia - Instruccion_page_2_cropped_words\
no.png
Saved cropped word 'ay' at Buendia - Instruccion_page_2_cropped_words\

ay.png
Saved cropped word 'muchu' at Buendia -
Instruccion_page_2_cropped_words\muchu.png
Saved cropped word 'pure-' at Buendia -
Instruccion_page_2_cropped_words\pure-.png
Saved cropped word 'za' at Buendia - Instruccion_page_2_cropped_words\
za.png
Saved cropped word 'de' at Buendia - Instruccion_page_2_cropped_words\
de.png
Saved cropped word 'conciencia,' at Buendia -
Instruccion_page_2_cropped_words\conciencia,.png
Saved cropped word 'Crezcan' at Buendia -
Instruccion_page_2_cropped_words\Crezcan.png
Saved cropped word 'en' at Buendia - Instruccion_page_2_cropped_words\
en.png
Saved cropped word 'vueftro' at Buendia -
Instruccion_page_2_cropped_words\vueftro.png
Saved cropped word 'fanto' at Buendia -
Instruccion_page_2_cropped_words\fanto.png
Saved cropped word 'temor,' at Buendia -
Instruccion_page_2_cropped_words\temor,.png
Saved cropped word 'y' at Buendia - Instruccion_page_2_cropped_words\
y.png
Saved cropped word 'amor,' at Buendia -
Instruccion_page_2_cropped_words\amor,.png
Saved cropped word 'co-' at Buendia -
Instruccion_page_2_cropped_words\co-.png
Saved cropped word 'como' at Buendia -
Instruccion_page_2_cropped_words\como.png
Saved cropped word 'en' at Buendia - Instruccion_page_2_cropped_words\
en.png
Saved cropped word 'los' at Buendia -
Instruccion_page_2_cropped_words\los.png
Saved cropped word 'afios,y' at Buendia -
Instruccion_page_2_cropped_words\afios,y.png
Saved cropped word 'mucho' at Buendia -
Instruccion_page_2_cropped_words\mucho.png
Saved cropped word 'mas.' at Buendia -
Instruccion_page_2_cropped_words\mas..png
Saved cropped word 'Adelantenfe' at Buendia -
Instruccion_page_2_cropped_words\Adelantenfe.png
Saved cropped word 'en' at Buendia - Instruccion_page_2_cropped_words\
en.png
Saved cropped word 'la' at Buendia - Instruccion_page_2_cropped_words\
la.png
Saved cropped word 'virtud' at Buendia -
Instruccion_page_2_cropped_words\virtud.png
Saved cropped word ', ' at Buendia -
Instruccion_page_2_cropped_words\,.png

Saved cropped word 'como' at Buendia -
Instruccion_page_2_cropped_words\como.png
Saved cropped word 'en' at Buendia - Instruccion_page_2_cropped_words\
en.png
Saved cropped word 'Jas' at Buendia -
Instruccion_page_2_cropped_words\Jas.png
Saved cropped word 'letras' at Buendia -
Instruccion_page_2_cropped_words\letras.png
Saved cropped word ',' at Buendia -
Instruccion_page_2_cropped_words\,.png
Saved cropped word 'y' at Buendia - Instruccion_page_2_cropped_words\
y.png
Saved cropped word 'mucho' at Buendia -
Instruccion_page_2_cropped_words\mucho.png
Saved cropped word 'mas' at Buendia -
Instruccion_page_2_cropped_words\mas.png
Saved cropped word ';' at Buendia -
Instruccion_page_2_cropped_words\;.png
Saved cropped word 'haf-' at Buendia -
Instruccion_page_2_cropped_words\haf-.png
Saved cropped word 'ta' at Buendia - Instruccion_page_2_cropped_words\
ta.png
Saved cropped word 'que' at Buendia -
Instruccion_page_2_cropped_words\que.png
Saved cropped word 'Ileguen' at Buendia -
Instruccion_page_2_cropped_words\Ileguen.png
Saved cropped word ',' at Buendia -
Instruccion_page_2_cropped_words\,.png
Saved cropped word 'por' at Buendia -
Instruccion_page_2_cropped_words\por.png
Saved cropped word 'vueftra' at Buendia -
Instruccion_page_2_cropped_words\vueftra.png
Saved cropped word 'imi-' at Buendia -
Instruccion_page_2_cropped_words\imi-.png
Saved cropped word 'tacion' at Buendia -
Instruccion_page_2_cropped_words\tacion.png
Saved cropped word ',' at Buendia -
Instruccion_page_2_cropped_words\,.png
Saved cropped word 'a' at Buendia - Instruccion_page_2_cropped_words\
a.png
Saved cropped word 'fer' at Buendia -
Instruccion_page_2_cropped_words\fer.png
Saved cropped word 'varones' at Buendia -
Instruccion_page_2_cropped_words\varones.png
Saved cropped word 'perfeós,' at Buendia -
Instruccion_page_2_cropped_words\perfeós,.png
Saved cropped word 'y' at Buendia - Instruccion_page_2_cropped_words\
y.png
Saved cropped word 'confumados' at Buendia -

Instruccion_page_2_cropped_words\confumados.png
Saved cropped word ',' at Buendia -
Instruccion_page_2_cropped_words\,.png
Saved cropped word 'agradables' at Buendia -
Instruccion_page_2_cropped_words\agradables.png
Saved cropped word 'vueftros' at Buendia -
Instruccion_page_2_cropped_words\vueftros.png
Saved cropped word 'ojos' at Buendia -
Instruccion_page_2_cropped_words\ojos.png
Saved cropped word ',' at Buendia -
Instruccion_page_2_cropped_words\,.png
Saved cropped word 'y' at Buendia - Instruccion_page_2_cropped_words\
y.png
Saved cropped word 'provechofos' at Buendia -
Instruccion_page_2_cropped_words\provechofos.png
Saved cropped word 'la' at Buendia - Instruccion_page_2_cropped_words\
la.png
Saved cropped word 'Republica' at Buendia -
Instruccion_page_2_cropped_words\Republica.png
Saved cropped word ',' at Buendia -
Instruccion_page_2_cropped_words\,.png
Saved cropped word 'que' at Buendia -
Instruccion_page_2_cropped_words\que.png
Saved cropped word 'libra' at Buendia -
Instruccion_page_2_cropped_words\libra.png
Saved cropped word 'cafi' at Buendia -
Instruccion_page_2_cropped_words\cafi.png
Saved cropped word 'to-' at Buendia -
Instruccion_page_2_cropped_words\to-.png
Saved cropped word 'da' at Buendia - Instruccion_page_2_cropped_words\
da.png
Saved cropped word 'fu,' at Buendia -
Instruccion_page_2_cropped_words\fu,.png
Saved cropped word 'felicidad' at Buendia -
Instruccion_page_2_cropped_words\felicidad.png
Saved cropped word 'en' at Buendia - Instruccion_page_2_cropped_words\
en.png
Saved cropped word 'la' at Buendia - Instruccion_page_2_cropped_words\
la.png
Saved cropped word 'acertada' at Buendia -
Instruccion_page_2_cropped_words\acertada.png
Saved cropped word 'he' at Buendia - Instruccion_page_2_cropped_words\
he.png
Saved cropped word 'crian:' at Buendia -
Instruccion_page_2_cropped_words\crian:.png
Cropping complete. All word images are saved in the
'cropped_words_pg_1' directory.
Saved: !.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\!.png

Skipped: ,.png (No suitable match found)
Saved: ,donde.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\donde.png
Skipped: -.png (No suitable match found)
Skipped: 2.png (No suitable match found)
Skipped: 4.png (No suitable match found)
Skipped: 6.png (No suitable match found)
Saved: ;.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\;.png
Saved: a.png -> Buendia - Instruccion_page_2_cropped_words_corrected\
a.png
Saved: abrazos..png -> Buendia -
Instruccion_page_2_cropped_words_corrected\abrazos.png
Saved: acertada.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\acertada.png
Saved: Adelantenfe.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Adelantense.png
Saved: Admirable.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Admirable.png
Skipped: afios,y.png (No suitable match found)
Saved: afsiftÃ©cia.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\assistencia.png
Skipped: ag.png (No suitable match found)
Saved: agradables.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\agradables.png
Saved: amor,.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\amor,.png
Saved: amor.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\amor.png
Saved: aparta.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\aparta.png
Saved: ay.png -> Buendia - Instruccion_page_2_cropped_words_corrected\
ay.png
Saved: bendecirles,.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\bendecirles,.png
Saved: bendiciones.,.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\bendiciones,.png
Saved: Benignidad.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Benignidad,.png
Saved: ca-.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\can-.png
Skipped duplicate: caf-.png (Already assigned to can-.png)
Saved: cafi.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\casi.png
Skipped duplicate: can-.png (Already assigned to can-.png)
Saved: celeftial.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\celestial.png
Saved: chriftiana.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\christiana.png

Saved: Cielos.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Cielos,.png
Saved: cion.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\cion.png
Saved: co-.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\co-.png
Saved: como.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\como.png
Saved: comunicais.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\comunicais.png
Saved: con.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\con.png
Saved: conciencia,.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\conciencia.png
Saved: condicion.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\condicion.png
Saved: confumados.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\consumados,.png
Saved: contento.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\contento.png
Saved: Crezcan.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Crezcan.png
Skipped duplicate: crian (Already assigned to cion.png)
Saved: da.png -> Buendia - Instruccion_page_2_cropped_words_corrected\
da.png
Saved: dad..png -> Buendia -
Instruccion_page_2_cropped_words_corrected\dad.png
Saved: dais.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\dais.png
Saved: de.png -> Buendia - Instruccion_page_2_cropped_words_corrected\
de.png
Saved: de=.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\de-.png
Saved: definir,y.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\definir,.png
Skipped: delâ€™.png (No suitable match found)
Saved: devora.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\devota.png
Saved: devotos.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\devotos.png
Saved: Dichofa.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Dichosa.png
Skipped: diffefio.png (No suitable match found)
Skipped: diffeiio.png (No suitable match found)
Saved: dignacion.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\dignacion,.png
Saved: dimiento.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\dimiento,.png
Saved: Dios.png -> Buendia -

Instruccion_page_2_cropped_words_corrected\Dios.png
Skipped duplicate: dis.png (Already assigned to dais.png)
Saved: divinas.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\vinas.png
Saved: dor.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\por.png
Saved: duria.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\duria.png
Skipped: duria.Vos.png (No suitable match found)
Skipped: dÃ©.png (No suitable match found)
Saved: echarles.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\echarles.png
Saved: edad,.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\edad,.png
Saved: edad.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\edad.png
Saved: efto.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\esto.png
Saved: el.png -> Buendia - Instruccion_page_2_cropped_words_corrected\
el.png
Saved: en.png -> Buendia - Instruccion_page_2_cropped_words_corrected\
en.png
Saved: enten-.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\enten-.png
Saved: entra.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\entra.png
Saved: es.png -> Buendia - Instruccion_page_2_cropped_words_corrected\
es.png
Saved: Eterno,.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Eterno,.png
Saved: exemplo,.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\exemplo,.png
Saved: explicarfe.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\explicarse.png
Saved: faber.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\saber,.png
Saved: fabi-.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Sabi-.png
Saved: fabios,.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\sabios,.png
Saved: fagrado.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\sagrado.png
Skipped: falÃ©n.png (No suitable match found)
Saved: fanto.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\santo.png
Saved: favo-.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\favo-.png
Skipped: fe.png (No suitable match found)
Saved: felicidad.png -> Buendia -

Instruccion_page_2_cropped_words_corrected\felizidad.png
Saved: feo.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\seo.png
Saved: fer.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\ser.png
Saved: fervorofamen-.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\fervorosamen-.png
Skipped: fi.png (No suitable match found)
Skipped: fingularsMaef=.png (No suitable match found)
Saved: fino.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\sino.png
Saved: foberana.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\soberana.png
Saved: folamente.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\solamente.png
Saved: fu,.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\fue.png
Skipped duplicate: fue.png (Already assigned to fue.png)
Saved: fueren.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\fueren.png
Saved: fus.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\sus.png
Saved: f{ua-.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\sua-.png
Saved: gion.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\gion.png
Saved: guftando.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\gustando.png
Saved: guro.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\guro.png
Saved: ha.png -> Buendia - Instruccion_page_2_cropped_words_corrected\
ha.png
Saved: haf-.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\has-.png
Skipped: he.png (No suitable match found)
Saved: hermanos.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\hermanos.png
Saved: hijos.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\hijos,.png
Saved: hu-.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\hu-.png
Skipped: Ia.png (No suitable match found)
Saved: ignorancia.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\ignorancia.png
Saved: Ileguen.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\lleguen,.png
Saved: imi-.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\imi-.png
Saved: indifpenfable.png -> Buendia -

Instruccion_page_2_cropped_words_corrected\indispensable.png
Saved: indignais.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\indignais.png
Saved: infinita.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\infinita.png
Saved: inocencia.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\inocencia,.png
Saved: Jas.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\mas.png
Saved: Jeru-.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Jeru-.png
Saved: la.png -> Buendia - Instruccion_page_2_cropped_words_corrected\
la.png
Skipped: laluz,y.png (No suitable match found)
Saved: leche.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\leche.png
Saved: les.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\les.png
Saved: letras.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\letras,.png
Saved: liberal.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\liberal.png
Saved: libra.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\libra.png
Saved: los.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\los.png
Saved: Manpre,.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\MADRE,.png
Skipped duplicate: mas..png (Already assigned to mas.png)
Skipped duplicate: mas.png (Already assigned to mas.png)
Saved: mayores,.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\mayores,.png
Saved: merecid.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\merecio.png
Saved: modettia,.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\modestia,.png
Saved: mucha.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\mucha.png
Saved: mucho.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\mucho.png
Saved: mudado.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\mudado.png
Saved: natural.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\natural.png
Saved: necefsi-.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\necessi-.png
Saved: Ni.png -> Buendia - Instruccion_page_2_cropped_words_corrected\
Ni.png
Saved: Nifio.png -> Buendia -

Instruccion_page_2_cropped_words_corrected\Niño.png
Saved: Nifios.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Niños.png
Saved: no.png -> Buendia - Instruccion_page_2_cropped_words_corrected\
no.png
Saved: norma.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\norma.png
Saved: obediencia.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\obediencia.png
Saved: oir,.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\oir,.png
Saved: ojos.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\ojos,.png
Saved: os.png -> Buendia - Instruccion_page_2_cropped_words_corrected\
os.png
Saved: Padres.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Padres.png
Saved: para.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\para.png
Saved: parece.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\parece.png
Saved: pecho.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\pecho.png
Saved: perfeA©os,.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\perfectos,.png
Saved: piedad.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\piedad.png
Skipped duplicate: por.png (Already assigned to por.png)
Saved: porten.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\porten.png
Saved: pregun-.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\pregun-.png
Saved: profeguid,.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\proseguid,.png
Saved: profeguid.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\proseguid.png
Saved: prometeis.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\prometeis.png
Saved: proponeis.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\proponeis.png
Saved: protec-.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\protec-.png
Saved: provechofos.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\provechosos.png
Saved: pudo.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\pudo.png
Saved: pues.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\pues.png
Saved: pure-.png -> Buendia -

Instruccion_page_2_cropped_words_corrected\pure-.png
Saved: que.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\que.png
Saved: quien.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\quien.png
Saved: recerles..png -> Buendia -
Instruccion_page_2_cropped_words_corrected\recerles.png
Skipped: revolver.Bien.png (No suitable match found)
Saved: regalados.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\regalados.png
Saved: Reli-.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Reli-.png
Saved: rendida;.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\rendida.png
Saved: Republica.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Republica,.png
Saved: Reyno..png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Reyno.png
Skipped: rifios.png (No suitable match found)
Skipped duplicate: Sabi-,.png (Already assigned to Sabi-.png)
Skipped duplicate: Sabi-.png (Already assigned to Sabi-.png)
Saved: Sean.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Sean.png
Saved: Seran.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Seran.png
Saved: ta.png -> Buendia - Instruccion_page_2_cropped_words_corrected\
ta.png
Saved: tacion.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\tacion,.png
Saved: tamibien.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\tambien.png
Saved: tan.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\tan.png
Saved: tar,.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\tar,.png
Saved: te.png -> Buendia - Instruccion_page_2_cropped_words_corrected\
te.png
Saved: temor,.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\temor,.png
Saved: Templos;la.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Templos.png
Saved: tienen.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\tienen.png
Saved: tierno,'y.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\tierno,.png
Skipped duplicate: tierno,.png (Already assigned to tierno,.png)
Saved: to-.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\to-.png
Saved: tos.png -> Buendia -

```

Instruccion_page_2_cropped_words_corrected\tos.png
Saved: tro.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\tro.png
Saved: unifteis.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\unisteis.png
Saved: varones.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\varones.png
Saved: vifsimos.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\vissimos.png
Saved: virtud.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\virtud,.png
Saved: Vos,.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Vos,.png
Skipped duplicate: Vos,.png (Already assigned to Vos,.png)
Saved: Vos.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Vos.png
Saved: vueftra.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\vuestra.png
Saved: vueftras.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\vuestras.png
Saved: Vueftro.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\Vuestro.png
Saved: vueftros.png -> Buendia -
Instruccion_page_2_cropped_words_corrected\vuestros.png
Skipped duplicate: vuettra.png (Already assigned to vuestra.png)
Saved: y.png -> Buendia - Instruccion_page_2_cropped_words_corrected\
Y.png
Saved: za.png -> Buendia - Instruccion_page_2_cropped_words_corrected\
za.png
Skipped: ~.png (No suitable match found)
Skipped: Â»y.png (No suitable match found)
Skipped: â€~de.png (No suitable match found)
Skipped: â€~mas.png (No suitable match found)
Saved: â€~mildad..png -> Buendia -
Instruccion_page_2_cropped_words_corrected\mildad.png
Skipped: â€™.png (No suitable match found)
Processing complete! All corrected images are saved in
'cropped_words_pg_1_corrected'.

```

Evaluate OCR Accuracy on Cropped Word Images Against Transcript Ground Truth

```

import os
import cv2
import pytesseract
import string

# Paths
word_images_dir = output_dir # Directory with cropped word images

```

```

ground_truth_file = transcript_file # Use the transcript file for ground truth

# Function to normalize words (lowercase, remove punctuation)
def normalize(word):
    return word.strip().translate(str.maketrans('', '', string.punctuation)).lower()

# Loading and filtering ground truth words
ground_truth_words = []
with open(ground_truth_file, "r", encoding="utf-8") as f:
    for line in f:
        normalized = normalize(line)
        if normalized and any(c.isalnum() for c in normalized): # Skip lines with only punctuation
            ground_truth_words.append(normalized)

ground_truth_set = set(ground_truth_words)

# Loading all .png images
image_files = sorted([f for f in os.listdir(word_images_dir) if f.lower().endswith(".png")])

predicted_words = []
matched_words = []

# OCR and evaluation loop
for image_file in image_files:
    img_path = os.path.join(word_images_dir, image_file)
    img = cv2.imread(img_path)

    if img is None:
        print(f"△ Could not read image: {image_file}. Skipping...")
        continue

    # Preprocessing
    gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
    gray = cv2.resize(gray, None, fx=2, fy=2, interpolation=cv2.INTER_CUBIC)
    _, thresh = cv2.threshold(gray, 0, 255, cv2.THRESH_BINARY + cv2.THRESH_OTSU)

    # OCR
    config = "--psm 7"
    ocr_text = pytesseract.image_to_string(thresh, config=config).strip()
    norm_pred = normalize(ocr_text)

    if norm_pred:
        predicted_words.append(norm_pred)

```

```

        if norm_pred in ground_truth_set:
            matched_words.append(norm_pred)

# Evaluation
total_predicted = len(predicted_words)
total_matched = len(matched_words)
accuracy = total_matched / total_predicted if total_predicted > 0 else 0

# Results
print(f"Total word images processed: {len(image_files)}")
print(f"Valid OCR predictions made : {total_predicted}")
print(f"Predicted words matched in transcript: {total_matched}")
print(f"Matching Accuracy: {accuracy * 100:.2f}%")

△ Could not read image: afsiftÃ©cia.png. Skipping...
△ Could not read image: delâ€™.png. Skipping...
△ Could not read image: dÃ©.png. Skipping...
△ Could not read image: falÃ©n.png. Skipping...
△ Could not read image: perfeÃ©os,.png. Skipping...
△ Could not read image: Â»y.png. Skipping...
△ Could not read image: â€™de.png. Skipping...
△ Could not read image: â€™mas.png. Skipping...
△ Could not read image: â€™mildad..png. Skipping...
△ Could not read image: â€™.png. Skipping...
Total word images processed: 209
Valid OCR predictions made : 193
Predicted words matched in transcript: 101
Matching Accuracy: 52.33%

```