Stan Development Team. 2024. *RStan: the R interface to Stan.*

Su, Hongjin, Weijia Shi, Jungo Kasai, Yizhong Wang, Yushi Hu, Mari Ostendorf, Wen-tau Yih, Noah A. Smith, Luke Zettlemoyer, and Tao Yu. 2023. One Embedder, Any Task: Instruction-Finetuned Text Embeddings. In *Findings of the Association for Computational Linguistics: ACL 2023,* edited by Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki, 1102–1121. Toronto, Canada: Association for Computational Linguistics, July. Accessed November 24, 2024. https://doi.org/10.18653/v1/2023.findings-acl.71.

Su, Jianlin, Jiarun Cao, Weijie Liu, and Yangyiwen Ou. 2021. *Whitening Sentence Representations for Better Semantics and Faster Retrieval,* arXiv:2103.15316, March. Accessed September 10, 2025. https://doi.org/10.48550/arXiv.2103.15316. arXiv: 2103.15316 [cs].

Swedenburg, Ted. 1995. *Memories of revolt: the 1936 - 1939 rebellion and the Palestinian national past.* Minneapolis, Minn: University of Minnesota Press.

Szojka, Zsofia A., Annabelle Nicol, and David La Rooy. 2020. Narrative coherence in multiple forensic interviews with child witnesses alleging physical and sexual abuse. *Applied Cognitive Psychology* 34, no. 5 (September): 943–960. ISSN: 0888-4080, 1099-0720, accessed November 24, 2024. https://doi.org/10.1002/acp.3673.

Tamari, Salim. 2008. *Mountain against the Sea: Essays on Palestinian Society and Culture.* University of California Press, November. ISBN: 978-0-520-94242-4.

Tamuz, Omer, Ce Liu, Serge Belongie, Ohad Shamir, and Adam Tauman Kalai. 2011. *Adaptively Learning the Crowd Kernel,* arXiv:1105.1033, June. Accessed September 28, 2025. https://doi.org/10.48550/arXiv.1105.1033. arXiv: 1105.1033 [cs].

Thurstone, L. L. 1927. A law of comparative judgment. *Psychological Review* (US) 34 (4): 273–286. ISSN: 1939-1471. https://doi.org/10.1037/h0070288.

Toivanen, Mari, and Bahar Baser. 2019. Remembering the Past in Diasporic Spaces: Kurdish Reflections on Genocide Memorialization for Anfal. *Genocide Studies International* 13, no. 1 (November): 10–33. ISSN: 2291-1847, 2291-1855, accessed November 26, 2024. https://doi.org/10.3138/gsi.13.1.02.

Transkriptor. 2024. *Transkriptor API documentation.* https://developer.transkriptor.com.

Vector embeddings - OpenAI API. https://platform.openai.com. Accessed October 1, 2025.

Voytiv, Sofiya. 2024. Diasporic group boundaries and solidarity in the making: collective memory in the anti-war protests in Sweden. *Ethnic and Racial Studies* 47, no. 2 (January): 391–410. ISSN: 0141-9870, 1466-4356, accessed November 26, 2024. https://doi.org/10.1080/01419870.2023.2261289.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data.* 2nd ed. Cambridge, MA, USA: MIT Press, October. ISBN: 978-0-262-23258-6.

Yahya, Abbad. 2017. Oral History and Dual Marginalization: *Jerusalem Quarterly.*

Zhou, Kaitlyn, Kawin Ethayarajh, Dallas Card, and Dan Jurafsky. 2022. Problems with cosine as a measure of embedding similarity for high frequency words. In *Proceedings of the 60th annual meeting of the association for computational linguistics (volume 2: short papers),* edited by Smaranda Muresan, Preslav Nakov, and Aline Villavicencio, 401–423. Dublin, Ireland: Association for Computational Linguistics, May. https://doi.org/10.18653/v1/2022.acl-short.45. https://aclanthology.org/2022.acl-short.45/.

Zhu, Lixing, Runcong Zhao, Lin Gui, and Yulan He. 2023. *Are NLP Models Good at Tracing Thoughts: An Overview of Narrative Understanding,* arXiv:2310.18783, October. Accessed October 21, 2024. https://doi.org/10.48550/arXiv.2310.18783. arXiv: 2310.18783.

Data Appendix

Extraction of POHA data

We extracted both audio files and corresponding metadata from the Palestinian Oral History Archive by scraping their website, located on the American University of Beirut's website at libraries.aub.edu.lb/poha/. Every record in POHA has its own dedicated ID, ranging from 4158 to 4887; each record's page is accessible via the URL `https://libraries.aub.edu.lb/poha/Reco`

For each interview page, we use the `BeautifulSoup` library (Richardson) to extract data from the HTML. Some of the fields (e.g., significant figures who appear in an interview) are variously present and absent; we flexibly add these to JSON representations of the metadata. Interview audio was downloaded from the "Interview Audio/Video" field present for most interviews. A minority of the interviews did not have available recordings; after this scraping process, during which we excluded those, we were left with 720 interviews that we were able to transcribe.

For analysis, we use a subset of these interviews. We unfortunately had to filter out a small set of interviews for which we were not able to extract the required metadata for our analysis (gender, place, place of origin, and place of residence). We also conducted further filtering of interviews that involved multiple participants from different places of origin, as this would have biased the results of our location analysis. The resulting set of interviews for analysis was 524.

Extracting Information from Bios

Our initial extracted metadata directly from POHA did not include specific fields for gender or place of residence. However, this information was sometimes also present within the metadata, just not in a structured format. When available, such information was often embedded within the *bio* field.

To extract gender and place of residence from the bios, which was essential for our study, we utilized a large language model. We compiled the bios in both Arabic and English and supplied them to `gpt-4o` OpenAI et al. 2024 using the following prompt:

```
From the following *interview description*
    + *metadata on names and places of
    origin* extract the following
    information in JSON format.

The input data has the following structure.

{{
  "en": {{"bio": "<bio in english>",
          interviewee: ["<name person 1>",
              "<name person 2>", ...],
          place_of_origin: ["<origin person
               1>", "<origin person 2>",
              ...]}}
  "ar": {{"bio": "<bio in arabic>",
          interviewee: ["<name person 1>",
              "<name person 2>", ...],
          place_of_origin: ["<origin person
               1>", "<origin person 2>",
```

```
            ...]}},
}}


The output data must have the following
    structure. Not all fields might be
    present; in that case, omit those

‘‘‘json
{{
  "en": {{  # for english data extraction
    "interview_date": ""
    "interview_location": ""
    "interviewees": [
      {{
        "name": "",
        "place_of_origin": "",
        "gender": "",
        "birth_date": "",
        "place_of_residence": "",
        "occupation": "",
        "occupation_entity": ""
      }},
    ]
  }},
  "ar": {{  # for arabic data extraction
    "interview_date": ""
    "interview_location": ""
    "interviewees": [
      {{
        "name": "",
        "place_of_origin": "",
        "gender": "",
        "birth_date": "",
        "place_of_residence": "",
        "occupation": "",
        "occupation_entity": ""
      }},
    ]
  }}
}}
‘‘‘

Notes:
- The interview might have more than one
    interviewee.
- The description is given in English and
    in Arabic, extract the data for both
    versions, and make sure the order of
    interviewees is the same for both
    versions
- Output the ‘name‘ as it appears in the ‘
    interviewee‘ list. If the name does not
    appear in the interview list, use the
    name as it appears on the bio
- Output the ‘place_of_origin‘ as it
    appears in the ‘place_or_origin‘ list.
    If the place does not appear in the
    place_of_origin list, use the place as
    it appears on the bio
- Use date format %Y-%m-%d (eg. 2024-08-25)
     if possible, if not, use %Y (eg. 2024)
- Be careful with keeping the correct
    transliteration of Arabic names and
    places

----------------------
Interview description:

{{
  "en": {{"bio": "{data["en_bio"]}",
           interviewee: {data["
               en_interviewee"]},
           place_of_origin: {data["
               en_place_of_origin"]}}}
  "ar": {{"bio": "{data["ar_bio"]}",
           interviewee: {data["
               ar_interviewee"]},
           place_of_origin: {data["
               ar_place_of_origin"]}}}
}}
```

Listing 1. Prompt to extract structured information from the bios

Note that in addition to the bios, we provided the interviewees' names and places of origin as recorded in the metadata. This ensured that the model used consistent names, simplifying the process of linking all data together.

Curating Places of Residence

Initially, our metadata extraction did not include direct information about the place of residence for each interviewee. While the bios sometimes contained additional details beyond the specified metadataincluding place of residencethis was not always the case. As explained in the previous section, we used `gpt-4o` to extract structured data from the bios.

After this process, many interviews still lacked specified places of residence in the bios, necessitating an alternative extraction method.

In our second approach, we focused on interviews for which no metadata was available. Arabic-speaking team members manually read the interview transcripts to identify the interviewees' places of residence. This task was inherently time-consuming, requiring approximately 30 minutes per interview for a native speaker due to the length of the interviews and the challenge of obtaining full context.

To expedite this process, we utilized `gpt-4o` with the following prompt:

```
From the following excerpt of an interview,
     extract the place of residence of the
    interviewee if such place is mentioned.

Notes:
- Do not confuse the place of residence
    with the place of origin of the
    interviewee.
- Attempt to identify the place where the
    interviewee is living at the time of the
     interview.
- It is very unlikely that the place of
    residence is within Palestine, as these
    are stories of exiled refugees.
```

```
Respond with the place of residence of the
    interviewee, the timestamp, and the line
     of text where you found it.
Also, add a short explanation of why the
    place is likely to be the place of
    residence.
If no place of residence can be reasonably
    identified, add a short explanation of
    why.

Text:
--------
speaker_id; timestamp; text
{text here}
```

Listing 2. Prompt to Extract Place of Residence from Interview Excerpt

Here, the text is formatted as specified in the prompt: `speaker_id; timestamp; text`.

Given that the interviews often exceeded the model's context window, we partitioned them by their table of contents and submitted each section individually. After processing a few interviews, our team members gained insights into which sections were more likely to mention the place of residence. Combining these insights with the model's suggestions, we efficiently processed approximately 250 interviews to obtain places of residence.

Out of the interviews that initially lacked residence information, we successfully identified the place of residence for 207 interviews. This left us with 52 interviews where the place of residence remained unidentified or unclear.

Deduplication of Places

We frequently encountered variations in how places of origin and places of residence were referenced within the metadata and bios. For example, a single refugee camp might be referred to in up to eight different ways due to the addition or omission of words, or the inclusion of the area or country name. This inconsistency posed challenges for grouping and matching interviews based on places of origin and residence.

To resolve this issue, we performed a deduplication process using the `RecordLinker` library (Farley and Gutman 2020). This tool allowed us to match locations by comparing both the Arabic and English versions of the names, utilizing edit distance metrics for each language. The process clustered records that likely referred to the same entity. We chose the most prevalent name in the data as the representative name for each cluster, effectively minimizing the impact of misspellings and variations by standardizing the names across the dataset.

After this process, we reduced the number of unique places to approximately 150 different places of origin and 100 different places of residence. We then manually curated these lists to finalize the mappings, ensuring accurate and consistent duplication of the interview information.

Transcriptions and validation

Filtering out interviews about folktales, we are left with 724 interviews. This still comprises about a thousand hours of audio and video. We use Transkriptor, a commercial transcription tool, to transcribe the interviews (Transkriptor 2024). Transcriptor supports transcription optimized for specific dialects of Arabic; this is important for the fidelity of our transcriptions of POHA, given that many of the interviews are in Levantine Arabic rather than MSA. While we explored using smaller Whisper (Radford et al. 2022) models, the large volume of interviews and the specificity of the data led to low-quality and slowly obtained results.

Two Arabic-speaking team members conducted a qualitative assessment of the interviews to validate the transcriptions. This process involved manually reviewing a subset of the transcriptions to ensure accuracy and consistency, particularly for dialect-specific nuances.

The translator effectively captured the local Palestinian dialect, without censoring words or omitting context, achieving 90% accuracy. However, challenges arose from factors such as the quality of the interviews and their structural design. In some cases, the transcription tool struggled to recognize sentence boundaries and properly connect the text, requiring careful contextual interpretation without re-listening to the audio. Additionally, there were minor issues with reading the local dialect, as the transcription tool spelled words differently than how native speakers would naturally express them.

Interview samples

For both interviews, only excerpts corresponding to the theme Exile, Expulsion and Displacement (filtered by TOC). The summary and key facts were generated via the summarization pipeline.

Interview 4634

Interview with Ahmad Salih Ismail, male, born in 1932 and part of the community in 'Amqa at the time of his expulsion from Palestine, where he worked in a fire match factory. At the time of the interview he resides as refugee in Nahr al-Barid Palestinian Refugee Camp, Lebanon.

Summary

تحمل تجربة المُحاوَر تفاصيل تهجير قسري بدءًا بحاصرة قريتهم ومغادرتها تحت وعود مضلّلة بالعودة، وسط غياب الدعم وتسليح كافٍ (q1,q2,q3). عايش المعارك المباشرة وفقد قريباً له أثناء القصف، ما زاد مشاعر الفقد والخوف (q4). مسيرة التهجير تضمنت تنقلات بين قرى ومدن، حيث عاش الإذلال والتفتيش والانتظار تحت لهيب الشمس (q5). تغيّرت حياته بالكامل، من عمله ومعارفه في فلسطين، إلى اللجوء والعمل الشاق في لبنان (q6). سنوات الإقامة في المخيم كانت مليئة بالمعاناة، حيث عاش في الشوادر والبركيات وسط ظروف غير إنسانية خاصة خلال الشتاء والعواصف (q7,q8). احتفظ المُحاوَر ومجتمعه برفض مشاريع التوطين والدمج حفاظاً على هويتهم، وعبّر بوضوح عن استمرار معاناتهم كلاجئين رغم مرور السنوات (q9,q10). ورغم كل القسوة، حملت الرحلة لحظات من التضامن والإحسان البسيط بين الناس (q11). تُظلّل هذه التجربة المستمرة شعور بالحزن والضياع، حيث تتداخل الذكريات القوية عن الوطن مع الواقع القاسي للمنفى.

The interviewees experience contains detailed accounts of forced displacement, beginning with the siege of their village and its departure under misleading promises of return, amid a lack of support and sufficient weapons (q1, q2, q3). They lived through direct battles and lost a relative during the shelling, which deepened feelings of loss and fear (q4). The displacement journey included movement between villages and cities, during which they experienced humiliation, searches, and long waits under the burning sun (q5). Their life changed completelyfrom work and acquaintances in Palestine to exile and hard labor in Lebanon (q6). The years spent in the camp were marked by suffering, as they lived in tents and makeshift shelters under inhumane conditions, especially during winter storms (q7, q8). The interviewee and their community consistently rejected resettlement and integration projects to preserve their identity, clearly expressing the ongoing suffering of refugee life despite the passage of years (q9, q10). Yet amid the harshness, the journey also included moments of solidarity and simple acts of kindness among people (q11). This ongoing experience is overshadowed by a deep sense of grief and displacement, where powerful memories of the homeland intertwine with the harsh reality of exile.

Key Facts

q1. خروج العائلة من القرية جاء نتيجة حصار واقتحام مسلح ووعود كاذبة بإمكانية العودة السريعة.
The family left the village after it was besieged and stormed by armed forces, and they were given false promises that they would be able to return soon.

q2. شهد المُحاوَر قصف الطيران على مناطق سكن أقاربه وشعر بالخوف لفقدانهم أثناء الهجوم.
The interviewee witnessed airstrikes on residential areas where his relatives live and felt afraid of losing them during the attack.

q3. انتقلت العائلة عبر عدة قرى ومدن وتعرضت لصعوبات أثناء الطريق وعمليات تفتيش وإذلال.
The family moved through several villages and towns, facing hardships along the way, including searches and humiliation.

q4. تغيرت حياة المُحاوَر بشكل جذري من حياة طبيعية وعمل إلى حالة لجوء في لبنان.
The interviewees life changed drastically, from a normal life and job to becoming a refugee in Lebanon.

q5. عانى المُحاوَر سنوات من العيش في مساكن مؤقتة مثل الشوادر والبركيات، وما رافقها من ظروف غير إنسانية، خاصة بفصل الشتاء.
The interviewee spent years living in temporary shelters like tents and makeshift huts, under inhumane conditions, especially during the winter.

q6. هناك رفض شعبي لمحاولات تثبيت اللاجئين من قبل المؤسسات الدولية واعتبار ذلك تهديد للهوية والانتماء.
There is popular resistance to efforts by international organizations to settle refugees permanently, as it is seen as a threat to identity and belonging.

q7. شعور دائم بالتهجير واللجوء مع غياب شعور الاستقرار، رغم مرور السنين.
A constant sense of displacement and exile, with no feeling of stability even after many years.

q8. تجربة التهجير رافقتها لحظات من التضامن الإنساني في ظروف القسوة.
The experience of displacement was accompanied by moments of human solidarity amid harsh conditions.

Interview 4634 — Amqa, Nahr al-Bared Camp, Lebanon

1 إسة بدك تقول لي كيف طلعتوا من القرية؟
So tell me, how did you leave the village?

2 ما قلنا لك كيف طلعنا؟ طوقوا البلد من التلات؟ جهات غربا من إبني ومن شمال وأدركوا الباب الشرقي مفتوح فتاكة رشاشة، يعني فقدنا كبس فرصة صارت الناس تطلع من البلد بإعتبار إنه بعد يومين ثلاثة من العرب شو يقولوا لنا إنه بعد يومين ثلاثة بعد أسبوع بترجعوا.
Didnt we already tell you how we got out? They surrounded the town from three sides: west, from the direction of my sons place, and from the north. They saw that the east gate was open, so they came in with machine guns. We lost any chance to stay, and people started leaving the town. They told us wed be back in two or three days even a week at most.

3 الملك عبد الله شو؟ قول ب15 أيار سيتكلم المدفع، كان التصريحات لإله يعني؟
What about King Abdullah? He said that on May 15 the cannons would speak. Those were his statements, right?!

4 والحمد لله ما شفنا منه هذا الشيء كله.
And thank God, we didnt see any of that happen.

5 بالزبط. لا والله.
Exactly. No, honestly.

6 بالظبط يعني بأي يوم لإسكان وقتها يمكن بهديك السنة، كان رمضان أول يوم برمضان أو ثاني يوم برمضان، طلعنا إحنا من البلد من أهمها يعني، فصارت المعركة.
Exactly. On that day, maybe that year, it was the first or second day of Ramadan when we left the village, basically. Thats when the fighting started.

7 يعني هالشباب اللي ما عندهم سلاح كل الليل؟
So, you mean the young men who didnt have weapons were up all night?

8 إنه مين طلع من عيلته؟ طلعتوا أنتم كلياتكم مع بعض الإخوة والأخوات؟
Who from your family left? Did all of you go together, brothers and sisters?

9 رشيحة قعدنا فترة نحنا طلعنا من البلد كلنا.
We stayed for a while as a group, then we all left the village together.

10 لما إجينا أكثر شي، قعدنا الفترة هذه.
When we first arrived, thats where we stayed for that period.

11 وصارت المعركة الأخيرة، واحتلوا اليهود ضربة الطائرات، أنا كنت ترشيح، أنا كنت قاعد بقهوة في الحارة الشرقية.
Then the final battle happened, and the area was taken after airstrikes. I was at that time sitting in a café in the eastern neighborhood.

12 إسم صاحب القهوة كان وقتها إسماعيل، بعدني ربكه. كان قاعد المختار. فهد شريح.
The café owner at the time was Ismail, I still remember him. The village headman was there too, Fahd Shurayh.

13 وشيء عشرة طنعشر واحد، ومنهم أخويا أبو نشيل أخوي الكبير، يعني وأنا قاعد معي أنا فراخ.
There were about ten or twelve people, including my older brother, Abu Nashil. I was sitting with them, just among the group.

14 عم بيقول المختار فات يشريع.
The village headman was saying that it was going to be resolved soon.

15 إنما الأسود قد غلبت الرئاب. كان طلع صوت الطيران، قام شو؟ قال له أخويا، أبو نجيب قال له والله يا مختار ما دام الطيران بالسمع ما في نجاح للعرب وضربة الطيار مع هذه الضربة، هن طلعوا الاختيارية، كنا نراحوا على بيت إي. جاي طابقين يعني قاعدوا بالطابق التحتاني، أنا قاعدة الناصرة.
Even the lions have been overpowered by the hyenas. The sound of aircraft had started overhead. My brother, Abu Najeeb, said to the village headman: As long as we can hear the planes, theres no victory for the Arabs. With that strike, old people started leaving. They went out toward the house, two floors, with the lower floor being where people were gathered, and I was sitting in the middle of it all. (Nazreath is mentioned as a word here, but not sure if it also means a location; the context is missing.)

16 مهبطة تحت أمطرة لا حالي. ولد. ا.
It was pouring rain, and I was by myself, just a kid.

17 اا بلش الطيران يقصف بالحارة الشرقي.
The planes started bombing the eastern neighborhood.

18 ضرب على البيت اللي هو فيه أخوي، أنا شفتهن سبع أظن نزلوا من الطيارة سبع.
They struck the house where my brother was. I saw them, about seven of them, I think, come down from the plane.

19 فراح البيض أنا قلت راح أخويا.
At that moment, I thought, my brother is gone.

20 إجينا تخبينا بذات الهيئة الكوليس كوليس إنجليزي لابسين بدلات بوليسي هني بس بلاد عربي يعني.
We came and hid with that unit, the police unit, dressed in English-style police uniforms, but in an Arab town.

21 فاتو على بيت فوتت معاها. لما ضربت الطيارة والشبابيك هاي اللي مثل هالشبابيك عنا. إذا صارت الشباك من هناك يضرب في الحيط من هون القزاز مثل الرصاص.
They came in and we went inside the house with them. When the planes started bombing, the windows, like the ones we have here, shattered. If a window was hit from that side, the glass would fly through the wall on the other side; the glass was like bullets.

22 فطلعوا. قالوا ما في إلنا غير حارة النسمة، لأنه حارة المسيحية.
So they left and said, we have no choice but to go to the Nasmeh neighborhood, because its a Christian neighborhood.

23 طلعنا على حارة المسيحية.
and so we went to the Christian neighborhood.

24 وإحنا وقتين سمعوا هن واحد عم يحكي بالبيت معهد طيران، هلا سكة.
And at that moment, we heard someone in the house saying something about the air force institute, like planes were on their way now.

25 الشباب، يعني أنا أولادكم فهمني الشباب، الزلمة يعني مسلحين ها وفاته عالبيت بس وقفوه أخذوه ووقفوا الطيران.
The young men, by that I mean the armed fighters, went into the house, but they were stopped, taken away, and the airstrikes were halted.

26–27 أنزلت أنا فقلت أخوي ما لقيته، فقالت أهلي لقيت إن طالعين. الحتة.
I went down and said I couldn't find my brother, and my family told me they had seen him leaving at the stop, the place.

28 لو.
If..

29 الحتة على الطريق، طلعنا على كفر اسماء بلد دروز.
We reached a stop along the road, then we headed out to Kafr Samaa, a Druze village.

30–32 إللي هو احتلو كل المنطقه أجو يوقفونا الكفر أسماء، أنا من الناس إللي قاعدو بالشمس من الصمغ، من الساعة 80. كل ساعة. قاعدين كانت عادية.
They had occupied the whole area and came to stop us at Kafr Asmaa. I was among the people sitting in the sun from the morning, around 8 o'clock hour after hour, sitting there as if it were normal.

33 دفعنا يوم الصبح، قالوا اللي يحب يطلع.
They moved us out in the morning and said, Whoever wants to leave can go.

34 يطلع يعني اليهود سمحوا للناس إنها تطلع إجا واحد من بلدنا إللي روحت نطلع أنا وياك، أنا هو واعي أوعى مني وأنا أصغر منه، وأهلي ولادي، شو بدي أصغر؟ قلت إن بقيوا هني يستغلوا.
Leave here means the zios allowed people to go out. A man from our village came and said, Lets get out, you and I. He was more mature than me, even though I was younger. My family and children? What was I supposed to do? I said we should stay here and wait.

35 بقي وأهلك بي كفر السنين بقيوا؟
Did your family, your father, stay in Kafr al-Sinin?

36 هن بكفر سميع ونظلقة الصبح.
They were in Kafr Sama, and we set out in the morning.

37 ونحنا طالعين. التقينا بأهل الدموع.
As we were leaving, we met the people of (Al-Dumu).

38 ومنهن بلاقي اللبان، أنا لا بعرفهم ولا شي.
And among them I found Laban. I didnt know them at all.

39–40 بس على الطريق إن أدى لواحد من، أه. من معار قوسوه الدروس مش اليمودة اللي كانوا عدنا على الطرقات، الد. الدروس، كانوا يشلحوا، يعني بعض الناس فقعدوا فتشوه ملائم مع مصاري قاموا فتشوهن بقوا 85 بعرف فلسطيني ربطهم على وقفوه على سفرة.
But on the road it led to one of them they were from the Druze of Maarr Qusuw, not the ones we used to see on the

roads. The Druze used to strip people down and search them. Some were detained and searched for their belongings and money; I remember about 85 Palestinians tied up and made to stand in a line on a table.

41 وقفوه على صخرة، وقوصوه بال بالرشاش الرشاش فيه صغير، قوسوه، ومعتزامن.
They made them stand on a rock and shot them with machine guns, small arms fire. They opened fire on them, at the same time.

42 ما عرفت اسمه ولا من النهار.
I didnt know his name, nor where he was from that day.

43 إيه طلعنا إحنا وإياهن، أنا طلعت لقيت دار أخوي، أبو نجيب كانوا قاعدين بعتها.
Yes, we left with them. When I went out, I found my brothers house, Abu Najeeb, and they were sitting inside it.

44 ب. بالجنوب يعني؟
Down south, you mean?

45 ودار أختي جوز أختي كمان مش هاي اللي أكبر، كانوا قاعدين هناك، أنا طلعت على غنيشة.
My sisters house, and my sisters husbands as well, not the older one, they were also staying there. I went out to Ghanisha.

46 طبع ا لأ؟ أنا مشقات كبيرة يعني من الجملة يمكن عجسر دير مراقنا لقينا في أربع مقتولين ومجلسين عالنهر يا مثلا.
Of course not. We went through very difficult hardships. When we reached Dir Marqana, we found four people killed and left sitting by the river, for example.

47 هناك كانوا يهودي ويهوديين، واقفين، واحد من شرهم.
There were zios men standing there, two of them, and one of them was from among their side, from their own group.

48 من شرقها، والبنت من قربها، كل ما إجا ناس يقولوا لهم
From the east side, and the girl was nearby. Every time people came, they would tell them

49–51 الجسر طويل. كان عليه لا يقل من 1500 شخص من طفل، وهذاك الطفل يسموه. الطفل أو الطفل اسمه الشبلة الكبير، الاختيار الكاهل اسمه.
The bridge is long. There were no fewer than 1,500 people on it, children included. One child was called 'Shibl al-Kabir' (the big cub), and the village headman the elder, the middle-aged mukhtar that was his name.

52 قعدوا يفتشونا ويسألوا هذا من أي بلد؟ أنا؟ قلت لهم أنا من أبقى.
They searched us and asked, Which village are you from? As for me, I said, Im from Baqa.

53 أنا حاسس أواعي بشروان و. ورطبته برقبتي.
I felt I was wearing a shroun (long robe/garment), and I tied its edge around my neck.

54 أم سب على البلد اليهودية؟ وكان بده يفتشني رجع لهون.
He swore at the Jewish town, and when he wanted to search me, he came back here.

55–57 سعد فرده جنبه. تسعة بربي أبو عجلين. بحطه بصدري.
Saad turned him to the side. Nine, I swear Abu Ajleen. I put it in my pocket, close to my chest.

58 بعد ما فتشني قلم. عادت عبيت شوك. مرار.
After they searched me, I got a pen. I went back and filled my pockets with thorns, painful memories, and bitterness.

59 ما حسيت فيه. والله ما حسيت.
I didnt feel it. I swear, I didnt feel it.

60 بعد ما فتشوا شوي، أنا صاروا يقولوا البلد الفلانية، بيروخ، البلد الفلانية. بيروح ايبه؟ ما جابوا سيرة البلد، ا، فأنت أنا ما رح تم قاعد لحتى يقعدوني، فزيت بين هالناس، ونسيت والله سبحانه وتعالى كتفهم يا شيخ.
After they finished searching me a bit, they started saying things like this village, send them there, that village, send them there. But they didnt mention my village. So I said to myself, Im not going to stay sitting here until they decide to detain me. I jumped into the crowd, and I honestly forgot everything, may God protect us, as you wish, Sheikh.

61 وصلنا على النور.
We arrived at Al-Nour.

62 المساء بعد ما وصلت الميش، أنا هناك مأساة كبيرة كانت للشعب الفلسطيني.
In the evening, after we arrived at Al-Mish, there was a great tragedy for the Palestinian people there.

63 .Love your mind

64 وإذا بدك تشربي المي، ما حدا يسقيكي بدبكة الناس، شربت من البركة اللي هي فيها قارش، وفيها الخنازير وفيها كل شي.
And if you wanted water, no one would give you any, youd drink from the pool where people and animals all gathered. I drank from a pond that had dirt, and even pigs and everything in it.

65 حتى الرعيم، أطموهن، أنا شفت شخص قام محرم من. من. من جنبه يمكن مبخض فيها طن لتبلها بالمي هذا؟
Even the wounded, they treated them. I saw someone whose side was injured, maybe badly, and they were trying to wash it with that water.

66 ما طلعت أنا وهذا الشاب قوة بقى برماشه. أنا رحت على عيتا عند أخوي الكبير يعني.
I didnt leave with them. I stayed with that young man for a while. Then I went to my older brother at Aita.

67 ثاني يوم الصبح إنزلي العرميش أسأل إنه حدا إجا من أهلي أو شي، وإلا واحد بيقولي أهلك إسة مرقو من هون، وطلعو على أنا.
The next morning, I went down to Al-Armesh to ask if anyone from my family had arrived or if there was any news. Then someone told me, Your family just passed through here and went on to Aita.

68 قانا بالجنوب.
Qana in the south

69 في الهند، أنا بس رب دين أبي الآن.
In India, Im just supporting my fathers religion now

70 مراقب مرق مقتنع بالوادي الحديث والوادي العيون نتيجة على حرية إجيت لقيت إنه على الطريق يعني قاعدين هن وعيلة من بلدنا، دار أبو حسين الدفير.
A checkpoint passed through, and I was convinced about the new valley and the valley of the springs. Then, because of the situation, when I arrived I found that on the road there were people from our village staying there, at Abu Hussein al-Dafirs house.

71 طبعا قعدنا هاديك الليلة بتنا وتاني يوم صاروا يقولوا إنه في عم بيودوا على سوريا، يعني لاجئين تروح على سوريا، بيوخدونا من.
We stayed that night and slept there. The next day they started saying that people were being taken to Syriathat refugees were being sent to Syria, that they would take us from

72 فطلعت أنا من طلعت أنا وأهلي هناك مع بعضنا، وصلنا ونحنا ماشيين على. أنا التقيت بشاب كنت أشتغل أنا وياه معمل الكبريت، هذا اليوم، قبل يومين كان إيه فلسطين؟ ثالث أنا وياه في لبنان يعني.
I left with my family, we all went together. While we were on the road, I met a young man I used to work with at the matches factory. That day, just a couple of days earlier, we had been in Palestine together. Now, we were meeting again in Lebanon.

73 طبعا أعطتنا بعضنا، وصرنا نمشي الجقتين، وأهلنا، إبقوا وصلنا أنا.
Of course, we met up with each other and continued walking in groups with our families. Eventually, we arrived, and I

74 ثاني يوم أجت البوستات، أخذتنا من هذا على وين؟ على الرشيدية، عصور.
The next day, the buses came and took us from therewhere to? To Rashidieh, in Tyre.

75–76 وأخوي هذا الكبير يا جماعة، نزلنا على. محطة صور، محطة تري.
My older brother guys we went down to the Tyre bus station.

77 Is your train. طلعنا for you.
At the Tyre station, we got off and then we headed out

78 بعنا بالترين يا ستنا. لو صرنا طرابلس.
We sold what we had at the train station, maam. Then we ended up in Tripoli

79–80 هاي، المحطة. المحطة هون والله.
Here, the station the station is right here, I swear.

81 بدي أحكي ضمير الإنسان، حتى العتال بطرابلس كان يحمل شوال خبز وينزل. يوزع للناس.
I want to talk about human conscience. Even a porter in Tripoli would carry a sack of bread and go down to distribute it to people.

82 أنا قعدنا تلات أيام.
We stayed for three days

83 إجا الحج الغندور. رحمة الله يعني.
Hajj Al-Ghandour came, may God have mercy on him (RIP).

84 فعلا زلمة كاملين، قال لا، لا، العثاني ولا الدولة، أنا اللي بدي أطعمني بعد إذنه مسجل الناس.
A fully grown man came and said, No, no, not the authorities or the state. I am the one who will feed you. Then, with permission, he started registering people.

85 بطلت تمرق I wish feedback brand and، the surrey of a حدا، سجل كل العائلات وصار شيء بأكل وطعميها، أنا أكلت منهن ثلاث أيام من هالفترة هاي.
No one was passing through anymore. He registered all the families and started providing food and feeding them. I ate from that for three days during that period.

86 طبعا رجعوا قالوا إنه أهل ترشيحه واللي ساكنين بالترشيح حيطلعوا بدهم يروحوا على سوريا، ليش؟ لأنه أهل ترشيح صار أكثر من هناك بسوريا.
Of course, they came back and said that the people of Rashidieh, those living in the campwould be sent to Syria. Why? Because there were already more people from Rashidieh in Syria.

87 إحنا طلعنا. طلعنا مع الطراشة.
We left. We went out with the buses.

88–89 المجذوب اللي كان ابنه وصل لدرجة عميد يعني بسوريا. ووو. حسين عثمان أبو زيد.
The Majdhoub, whose son eventually reached the rank of brigadier in Syria, and Hussein Othman Abu Zaid.

90 طبعا، خمس عائلة كنا، إنه ساجنين بيطرشيها، طلعنا معاه.
Of course, there were five families of us staying in Rashidieh, and we left with him.

91 أعدنا ست أيام لحمص رجعو وزعونا لوين بس؟ شو بدنا نحكيلك؟ أحكيلك عخمسة؟ يعني؟ شو اللي ثمنا بحمص؟ لقيت اللي ثمنا بحمص وبسكنة خالد بن الوليد؟ لو قلت لك الأمل هناك كان أكثر من بيت نمل، قولي أكثر شوي، يعني بيوت نمل يعني بيوت نمل وفايعة الأمن، يعني مجرد ما همت لبناني بتحسي إنه في شي؟
We stayed in Homs for six days. Then they came back and redistributed us, where exactly? What can I tell you? Should I tell you about five (places/areas)? What we experienced in Homs we found that even where we stayed in Homs, in the Khalid ibn al-Walid neighborhood, I cant even describe it fully. Let me put it this way: there were like ant houses, very crowded, extremely cramped living conditions. There was no real safety. Just hearing that someone was Lebanese made you feel like something was happening (like tension or risk).

92 بمرق عليكي كأنه العسل.
I pass by you like honey.

93 وإذا غيرتي أوعيتي وطلعت بلاقي عليهم أمن ما بنقدم.
And if you changed your clothes and went out, youd find security forces on themwe couldnt move forward.

94 فعدنا فطرة الناس ستي أم وزعون على القرآن؟
Then we went back. The people maybe they were distributing the Quran?

95–96 طلعت منطقتنا إسمها صابونية بلد إسمها الصابونية، تبعد بعمل قريتين يعني حدود العراق، مشيوا تيجي. من قبلة للشرق، من حمص.

The Saboonyeh area we went back to our region, called Al-Saboonyeh, a village about two villages away, near the Iraqi border. They were walking toward it from the east, from Homs.

97 في 7 ساعات من إيدي عن حمص السياره، أعدنا أقول هناك فتره شي ثلاث شهور إجا واحد من بلدنا.

Its about 7 hours by car from Homs. We stayed there for a period of about three months. Then someone from our village came.

98 من قريبنا يعني إجا يوخذ عمته كانت نايمة، عمتو بتكون خلته لأبو.

A relative of ours camehe came to take his aunt, who was sleeping. His aunt would be his fathers sister.

99 أنا طلعت معاي يعني اختيار.

I left by choice, I mean it was optional (not forced)

100 باجي، آخدها قال فيه الصليب الأحمر بدو يجي يسجل هون.

He said, Come, Ill take her, because the Red Cross is coming here to register people

101 كنت باجي، أخذكم قرنالهم. ماشي الحال والله، لأنه هناك ما فيهن لا تشتغلي ولا شي ثلاث.

I used to come and take you (and bring you along), I told them its okay. Honestly, it was fine, because there you couldnt work or do anything.

102 شهور.

Months.

103 ثلاث شهور.

Three months.

104 إيه، إجا الزلمة قال بلش الصليب الأحمر، وإذا تأخرتوا ما بيعود فيه، إنت تسجيل يللا كان وقفونا بدهن يرجعونا، قلنا هلق بنقعد بحمص بس نرجع للقرين بدنا ناجي.

Then the man came and said the Red Cross has started (registration), and if youre late, there wont be another chance. So we were stopped and they were going to send us back. We said, Well stay in Homs for now, but well go back to Al-Quraynwe have to go.

105 الله جابنا لهونيس.

God brought us here

106 جنب المسلح اللي عاملينه لسه بالوقت نفسه.

Next to the armed man they had set up, at the same time

107 قعدنا هيك فترة جابوا المخيم لهون.

We stayed like that for a while, then they brought the camp here

108 جاب وقت ما إجيتوا لهون، شو اشتغلتوا؟ شو كنتوا تشتغلوا؟

When you arrived here, what did you work? What kind of jobs did you do?

109 كنا نعدل؟

Did we used to do repairs / maintenance work?

110 نطلع. بحس أنا الورش نطلع بحس نطلع حجار ننزل بحجارير الوسخ بنظف شو بدو يشتغل؟ هذا شغلنا كان.

We used to go out. I felt like we were in workshops, wed go out, break stones, go down into sewage systems, clean the dirt, what kind of job is that supposed to be? That was our work.

111 بعدين تطور من الإعتال اللي تنكث البطون من تنكث البطون، صار محبب بالسيارات، بحثوا رمز.

Then it developed, from working in concrete/brick hauling jobsuntil he became interested in cars. it became more focused on cars. They started searching for symbols

112 لحد ما إجيتو على المخيم.

Until you came to the camp

113 لحتى وجينا المخيم المخيم، كانوا يشتغلوا هون ب1/4 ليرة، أنا كنت بنزل يجونا جابتنا الدولة.

Until we came to the camp, they used to work here for a quarter of a lira, and I used to go down the state brought us here.

114 قعدنا بطرابلس شي سنة أكثر من سنة، أكثر من سنة لغاية ما إجا المخيم لسنة ال51 كيف؟

We stayed in Tripoli for about a year, more than a year, until the camp came in 1951

115 إجيتوا عن المخيم. كيف إجوا ومين اللي أخدكم إجت.

You came to the camp. How did they come? Who took you? How did you get there?

116 الدولة بالوجه الصبح عملوا أوعيكم، بدناش بدنا ما بدناش حتى الدنيا يخلوا البراكيات اللي عملها الغندور للناس، لأنه الغندور، بعد ما طلعنا إحنا إعملهم براكيات.

In the morning, the authorities came and told you to get your things ready. They said they wouldnt leave the makeshift shelters that Al-Ghandour had built for the people. Because Al-Ghandour, after we left, was the one who built those shelters.

117 لأنه رأي الكاس خانة، شو بدك تقولي؟ يعني مثل إيه موقع عسكري بركسات للجيش كاين للجيش التركي القديم وعشرطة البحر؟

Because it was called Kaskhanawhat can you say? It was like a military site, military barracks for the army, originally for the old Ottoman Turkish army and for the police and navy.

118 الغندور؟ ما غني عمل براكيات قسمها يعني بمركيات أنا؟ شو يعني؟ إسلام؟ هذول الكتاب 1200 عائلة بدهم يفوتوا بقلب براس؟ كيف بدهم يناموا هدول؟

Al-Ghandour built makeshift shelters and divided them up. There were around 1,200 families who were supposed to go inside them

119 أنا ما رضي إلا السماء قبل ما فوتهم، يعني قعدوا هن أكثر من شهر.

I didnt agree until the place was cleared, before they let them in. They stayed there for more than a month.

120 وأعلى الناس فيها، ونحنا إجينا من الجملة لقينا بركية بعدنا إجينا.

And the people were crowded at the top, and when we came from the group, we found a shack left for us when we arrived

رحمة الله عليه. كان إنسان كثير.
May God have mercy on him, he was a very kind man.

كانوا عالعياد يذبح للناس يذبح يعني ويوزع للناس، طالما كنا قاعدين هناك كان بشكل يعني مثلا يمكن اللحم اللي يعطيها للناس أكثر إنه يشتروا الناس، اليوم كثير.
During the holidays, he used to slaughter animals for people and distribute the meat. While we were there, he would give out a lot of meat, more than people could even afford to buy.

الأعياد يبدء يعني بيوسع زي ما كان دجن كثير والله.
The holidays would pass like that; he used to be very generous. May God have mercy on him

ما حدا يشجع منكوا أبد ا وإنتوا عالطريقة السير.
Did no one encourage you while you were on the road?

في إيه؟ لأ، أنا والله ما رجعت، بس في ناس في واحد من بلدنا رجع اسم حسين المختار، رجع من حمص، أقول لك شو أسوي، كيف رجع؟ سأل قد إيه؟ كانوا الدخان؟ قالوا له 35 ليرة شوية.
No, I didnt go back, but some people did. One man from our village named Hussein Al-Mukhtar went back from Homs. Let me tell you what happened. How did he go back? He asked about the pricesabout cigarettes? They told him: 35 lira.

لكن نحن نوقف، قال عليه الطلاق غير يرجع لفلسطين وينشرها بشريده، إنه كيلو تقانه 35 ليرة والله، ورجع من حمص على فلسطين بعد الفلسطين، لا ما رجع أخد مرض وراح وبعده الفلسطيني.
But we stopped him. He said: I swear I will go back to Palestine and spread the news everywhere, if a kilo is 35 lira, by God. So he went back from Homs to Palestine. But after that, no, he didnt continuehe got sick and died, and the Palestinian story remained.

طيب وإنتو جايين صار مع حالات وفيات أو ورادات على الطريق؟
While you were coming, did any deaths or injuries happen with you on the road?

أبد ا؟
Not at all?

لأ. أنا ما أعرف يعني مع المجموعة اللي كنتو جايين.
No, I dont know. I mean, with the group you came with?

أبدا.
Never / Not at all.

طيب إجيتوا على المخيم، بدك تحكي لي كيف إجيتوا على المخيم؟ كيف إجت جابوها؟
Okay, you came to the camp. Can you tell me how you came to the camp? How did they bring you?

بالسيارات؟
By cars?

بالتفصيل.
In detail.

نزلنا وين الأرض إللي مخدها موعد إسا جنب المدرسة. حطونا فيها.
We went down to the land that had been designated for us, now next to the school, and they put us there

لما يجي الأيام. كان يسير شتى. كثير.
When those days came, a lot of different things would happen.

قعدنا أكمن يوم صار الشتاء، فاضت المايك ما يأرجوها، كان إلها كتر كمان يمكن زواعينا، أنت ما بعرف، كان إلها كتر في الأرض وهي جو فاا تلت الشوادر ما هي؟
We stayed for a few days, then winter came and the water overflowed and they couldnt control it. There was a lot of rain, maybe even storms. Im not sure exactly, but the water was going into the ground and flooding the tents.

الأواعي.
Clothes

متل.
Like what?

ما بتفصل؟
Doesnt it stop?

كانوا قاعدين بهاي المنطقة؟
Were they staying in this area?

وين الدوامة؟ لسا كانوا آمين منهم؟
Where was the storm/whirlpool? Were they still safe from them?

لما صار الشتاء وصار شغلات من هالشكل طلعوا الصلبابي على المنطقة اللي فوقها، قمنا نحنا طلعنا من تحت.
When winter came and things like that started happening, the Crusaders went up to the upper area, and we moved out from below

شو بدي أحكيلك؟ مأساة؟ الشوادر؟ مأساة الشوادر كان بالليل لما تقوم عواصل يتنزع الشابر أو يفلت الأوقات يمزعه وينامو تحت.
What can I tell you? The tents were a tragedy. The tents were a disaster. At night we would wake up, and because of the strong winds the tent would tear or come loose, sometimes it would rip completely, and we would still sleep underneath it

عملوا حراب بالأول للشهداء، أول ما حطينا ما أعرف لا أحراف ولا شي، بعدين صاروا يجيبوا فحصات النسوان والوكالة يعملوا له داير.
At the beginning they used to bury the martyrs quickly and roughly. When we first settled, I dont really know about tents or anything like that. Later they started bringing examinations for women, and the UN agency would handle arrangements for them.

انفلت ودي هالليل إن صارت عواصف، إشي يتمزق عليه الشهادة، إشي ينام تحت الشادر، إشي ينطق عند جيرانا بشادر ثانيه، مأساة الشوادر، مأساة، يعني تم ويعطونا شوادر لفترة.
That night things got out of control. When storms came, some tents would tear, some people would sleep under the same torn tarp, and some would go to neighbors in another tent. The tents were a tragedy, a real disaster. They would only give us tents for a limited period.

ده اللي هنا بزعشات. وبعدين صاروا. يعمل الناس هشاش؟ ويحطو عليها الشاذلي، يصبو، يحطو عليها قصب، ويحطو الشابر.
This here was just chaos, and then they started making temporary shelters for people they would put 'al-Shadali' over them, pour in materials, add reeds, and cover them with tarps.

149 بقينا بالمأساة للثلاثة لسنة ال58 المأساة بالشوادر الكبيرة كتير. يعني ما حدا بيقدر يوصفلك قد ما كانت مأساة للثمانية و50، أكثر الناس بدأت بالشوادر لغاية ال58 لما صارت الثورة.

We stayed in this tragedy until 1958. The tents were large and it was a huge disaster. No one can describe how bad things were in 1958. Most people were living in tents until 1958, until the revolution happened.

150 لبنان، فلم انصرف الثورة في لبنان؟ صارت الناس تقدر تعمل أوائل شطب طين وإشي يقلع جار من الأرض، ويعمره خشش.

Lebanon, after the revolution in Lebanon, people were able to start building basic houses out of mud, and they would take stones from the ground and build them with wood

151 وأملو زين كسارة الوكالة تطلع زين كسارة، ويحطو عليها زين برضه مأساة الزينكو غريب، هي عجيبة، يعني يمكن تلحقه إذا شفته بيوت زي كيف هي إذا قعدت تحتية، كأنك عم تنشوا بالنار.

And they would put zinc sheets from (UNRWA) used to distribute zinc as well. The zinc (roofs) were a tragedy, something strange. If you saw how those houses were, and if you sat under them, it felt like you were sitting under fire

152 طيب مين اللي نظمكم بالمخيم؟ بس؟ إجيتوا أو بس إجيوا نزلتوا على لبنان؟ ما في أي تنظيم؟

Okay, who organized you in the camp? Just that? You came, did you just arrive and settle in Lebanon? Was there no organization at all?

153 مشروع كلاب إني أول ما إجت الوكالة عندها بمشروع كلاب، كيف اللي بتداوي تعمر تعمر هذا المشروع؟ قرره هن عسبيلي، عمرو، هذا المخيم.

Project Kelab when the UNRWA first came to Project Kelab, how did they treat and develop that project? They decided it in my case and built this camp

154 فالناس اللي ما بدها واللي هي مش وطنيه بالأحرى، يعني لو إنه تعمر المخيم كان أخدنا إحنا هاي الأرض اللي إشتريناها كان أخذتها كلياتها الوكالة وعملت بيوت زي الناس، بس الناس صاروا يقولوا هذا مشروع كلاب ومشروع توطين وشغلات من هالشكل وهن كانوا كليتهم.

The people who didnt want it, and who were not patriotic, to be honest, if the camp had been properly developed, we would have taken the land we bought, and the UNRWA would have taken it all and built proper houses for people. But people started saying this was Project Kelab and a settlement project and things like that, these were all rumors.

155 من عملاء الوكالة وعملاء الدولة، من آخرها، رفضت الناس بالبداوي عمله.

Agents of the agency and agents of the state, in the end, the people in Beddawi refused to implement it.

156 ضبية عمله ايه النبطية عمله؟

They did it in Dbayeh, and yes, they did it in Nabatieh as well?

157 هون ما عملوا باقي المخيم على حاله يعني بدون تنظيم، يعني إنت أخذت شقة أرض عارف فيها هذا البيت، أنا و20 مشترك من إيجادنا هون ب40 ليرة، البيت تبعكم أخدته منه بالقوة بليرتين.

Here they didnt do anything, and the camp remained as it was, without any planning or organization. You took a plot of land and built this house on it. Twenty others from our group and I bought land here for forty lira. As for your house, you took it from him by force for two lira

158 مأساة الشوادر الكبيرة، كثير ما حدا بيقدر يوصف لك إياها قديش كانت نفسها.

The tragedy of living in the tents was immense. No one can truly describe just how terrible that suffering was

159–160 طيب شو كان موقف اللبنانيين؟ بس نزلتوا على رميش أهل رميش وغير كمان المناطق اللي نزلتوها؟ كل اللبنانيين بشكل عام؟

Okay, what was the stance of the Lebanese people? Starting from when you arrived in Rmeish how were the people of Rmeish, the people in the other areas where you stayed, and the Lebanese overall?

161–162 قرشين أو 5 قروش فلسطيني، حق الشوربة، وبتسقيكي قد ما بدهم أنا من الناس اللي شرب قرشين صبتلي بتم، بس بلا يعني. والكل يرفض حتى يسقيكي. بعدين الشيعة ما بتعرفي إنت إذا شربتي من إبريقهم؟ هم بيسؤوكي؟

Two or five Palestinian qirsh for soup, and they would serve you as much as they wanted I was among those who drank for two qirsh; they poured it into my mouth, but without much care. And everyone would even refuse to give you water. As for the Shia, you wouldn't even know whether you'd drunk from their water sources would they give you water?

163 يعني برضه ما بنقدر نلومهن، لأنه حتى شعبنا الفلسطيني لما البلد اللي تطلع ما تلاقي.

I mean, we also cant blame them, because even our Palestinian people, when they leave their country, they find nothing

164 شغلة منيحة، لأنه أنا بعرف إنه بالدير زته الزبل الدير القاسي، زته الزبل اللي بيرى من شانه، الناس ما تشرب.

Its a good thing. Because I know that in Deir, even the harsh garbage, the kind of garbage a person would be ashamed of, people dont drink (from it / near it)

165 هاي بفلسطين.

This is in Palestine.

166 ونحنا طالعين، يعني اللي قاعد بالدير، يعني والله نحنا قعدنا بترشيحك كانت معاملة.

And when we were leaving, the people in Deir honestly when we stayed in Tarshiha, the treatment was

167 كانت عدد من يعني عند واحد ابن المختار بيقولوله كامل الشريح.

There were some cases like one of the mukhtars sons, they used to call him Kamil Al-Shar.

168 كان زلمة محترم. الله يرحمه برحمته.

He was a respectable man. May God have mercy on him

169 أي جهات عسكرية، لبنانية أو فلسطينية؟

Any military sides, Lebanese or Palestinian?

170 وصاروا يطعموا الناس قتلوا من هون؟ يكتب تقرير ويبعته يلحقوه.

And they started feeding people were there reports of killings from here? They would write reports and send them and follow them up

171 طيب، جيش الإنقاذ إنه قال له أي دور هون؟
Okay, did the Arab Liberation Army (Jaysh al-Inqadh) have any role here?

172 بالجنوب بشوي، بعدين طرحوا على الجولان.
In the south a bit, then they withdrew to the Golan.

173 بالجون ما بعرف شو عادوا، يعني فترة بالجولان، وبعدين صاروا.
Then I dont know what happened to themso they stayed in the Golan for a while, and then they

174 يعني من هذا خليل عبده كان في الجولان، يمكن قاعد بيطلع شي سنة، وبعدين إجا على الكاس كان قبل ما طلعنا من الكاس خانة إجا تجوزن.
So Khalil Abdo was in the Golan, maybe he stayed there for a year, then he came to the Kaskhana. Before we left Kaskhana, he came and got married there

175 لساته الشعب الفلسطيني. لاجئ؟
Is the Palestinian people still considered refugees?

176 أول ما نزلت على لبنان.
As soon as I arrived in Lebanon

177 هون في المنطقة عنا، كانوا يقولوا إلهن ديون.
Here in our area they used to say they had debts.

178 أحكيلنا عن هذه القصة؟
Can you tell us about this story?

179 كانوا يقولوا إلنا ديون حتى في.
They used to tell us that we even had debts up to

180 من جملة اللي صار، لا في اللبان، كان في هيك واحد قاعد، وين الخان جنب الخان، كان قاعد هيك؟ اسمه ناظم ناظم بيك المرعي، فكان هذا لا في يطلع هو وبعض اختية يروح وهيك على شط البحر، لأنه كان في طريق هناك يعني شارع اسفلت يعني بس يشوفوني يفوتوا لجوا على الخارج يفوتوا على بيوتهم، فيوم نادى علينا لا في ولا بيقولوا يهدوا اللي يحكوا مثلنا.
Among the things that happened, there was a man in Lebanon sitting near the khan, his name was Nizam Bey al-Murri. He used to go out with some of the mukhtars to the seashore, because there was a paved road there. They used to see us passing by and going back into our homes. One time he called out to us and said they should silence people who talk like us.

181 قال له تعال، تعال، يا ريت. قال له شو؟ قال له؟ تفضل قديه تتابع؟
He told him: come, come, I wish you would

182 إحنا أوعى منكم 1,000,000 درجة بس ربنا كتب علينا، صار عاد صحبي هو ي صار صحبه يعني ويروح ويجي ويروح؟
We are a million times more aware than you, but God has decreed our fate. He became my friend, they became friends, going back and forth, coming and going

183–184 طيب صار لك أي إخوة بعد الهجرة؟ بعد يجي ال48؟ ولاد أو بنانا؟ لا ما في؟
Okay, did you have any siblings after the migration, after 1948? Children, daughters? None?

185 بس ولاد إخوة، كثير لا.
Only nephews? Many? No?

186 بدي إخوة عشان نعرف أسمائهم.
I want to know the brothers so we can learn their names

Interview 4637

Summary

تروي تجربة المُحاوَر مسار التهجير خلال النكبة، حيث داهمهم الخوف والرعب بفعل الهجمات والدعم الخارجي للمهاجمين، فأُجبرت العائلات على الهرب الجماعي حفاظًا على كرامتهم (q1,q2). بعد النزوح، عاشوا مرارة الحياة المؤقتة بالخيم والشوادر، وسط ألم وذل مستمر وصعوبة تأمين سبل العيش (q3,q5,q6,q8). رغم مساعدات الجهة الإغاثية بتوفير مأوى وطعام، ظلت المعاناة ولم تُعوض هذه المساعدات عن ويلات المنفى (q4,q8). في مكان اللجوء، واجه معاملة قاسية ومهينة من بعض السكان مما عمّق الشعور بالحرمان (q7). ما زال المُحاوَر يشعر بألم الفراق القسري عن الوطن وأن خروجه منه كان رغماً عنه (q10). يتمسك بقوة بعادات وتقاليد أجداده رغم الشتات، يراها ركيزة هويته (q9,q12). يتغلغل فيه حلم العودة إلى بيته وأرضه، رافضًا حياة المنفى الناقصة ويتمنى أن يختم عمره في الوطن (q11).

The interviewee recounts the displacement journey during the Nakba, where fear and terror spread due to attacks and external support for the attackers, forcing families into mass flight to preserve their dignity (q1,q2). After displacement, they endured the hardship of temporary life in tents and tarps, with ongoing pain, humiliation, and difficulty securing livelihoods (q3,q5,q6,q8). Despite relief assistance providing shelter and food, suffering continued and these aids did not compensate for the hardships of exile (q4,q8). In the place of refuge, they faced harsh and humiliating treatment from some locals, deepening their sense of deprivation (q7). The interviewee still feels the pain of forced separation from their homeland, believing they left against their will (q10). They strongly hold onto the traditions and customs of their ancestors despite displacement, seeing them as a core part of identity (q9,q12). A deep desire for return to their home and land persists, rejecting an incomplete life in exile and wishing to end their life in their homeland (q11).

Key Facts

q1. الشخص المُحاوَر وعائلته وجماعته اضطروا للهروب من مناطقهم خوفا على حياتهم وكرامتهم بسبب ترويع مجموعة مسلحة ودعم من جهة خارجية.
The interviewee, his family, and his community were forced to flee their areas out of fear for their lives and dignity, due to intimidation by an armed group and support from an external party.

q2. الحياة بعد التهجير تميزت بالحرمان الشديد والمعاناة المتواصلة تحت الخيم والشوادر، حيث كان الناس مجبورين يبحثوا عن قوت يومهم في ظروف مذلة.
Life after displacement was marked by severe deprivation and ongoing suffering under tents and tarps, where people were forced to search for their daily sustenance in humiliating conditions.

q3. اللاجئون تلقوا دعما من جهة إغاثة تقدم لهم مأوي مؤقت ومواد أساسية، لكنها لم تعوض قسوة المنفى ولا حلت صعوباته.
The refugees received support from a relief agency that provided them with temporary shelter and basic supplies, but it did not compensate for the harshness of exile nor solve its difficulties

q4. الشخص يشعر بمرارة من معاملة بعض المجتمعات في مكان اللجوء، ووجدها مهينة وغير إنسانية.
The person feels bitterness from the treatment by some communities in the place of refuge and finds it humiliating and inhumane

رغم قسوة التهجير والحرمان، الشخص متمسك بقوة بعادات الأجداد وتقاليد q5.
الضيافة الأصيلة كرمز للهوية والانتماء.
Despite the harshness of displacement and deprivation, the person strongly holds onto ancestral customs and authentic traditions of hospitality as a symbol of identity and belonging.

الشخص يحمل ألم فادح على فراق الوطن، ويعبّر عن الحلم الدائم بالعودة ورفضه q6.
لحياة المنفى المنقوصة.
The person carries deep pain over separation from the homeland and expresses a lasting dream of return and rejection of an incomplete life in exile.

Interview 4637 — Amqa, Nahr al-Bared Camp, Lebanon

1 اليهود إلى عمق.
The Jews into the depth

2 أجد أهل الكات وأهل الشيخ دهود وأهل دنون وأهل الغبش كلياتها على عمه.
find the people of Al-Katt, the people of Sheikh Dahoud, the people of Dannoun, and the people of Al-Ghabashall of them are on the same side / under one aunt (meaning: same group/lineage context

3 هارمي من اليهود.
A Jew is a nomad / from the Jews.

4 وعم أشرح ودخل عليها رويب كمان نفس الشي مثل ما داخل الرعب على أهل الكوكاد وأهل القطعية وأهل.
And Im explaining and Ruwaib also entered it the same way, like how fear entered the people of al-Kokad and the people of al-Qatia and the people of

5–6 نتيجتها من معر؟ وقوع الكارب، بالنسبة.
What was the result, from Ma'arr? The disaster came about, regarding

7 إشتراك هاي، هاي الضيعة كلها، والكوكات والغبشية، والشيخ داوود كانوا مشتركين في هذه المعركة، دخل علينا.
All the people of this village participated, and al-Kokat, al-Ghabshiya, and Sheikh Dawood were all involved in this battle, and it entered upon us

8–9 إيه الإنجليز، دعم اليهود، الدعم ومدن بأسلحة إيه يعني؟ حديثي تختلف عن أسلحتنا؟
Yes, the British supported the Jews with weapons and assistance their weapons were different from ours.

10 سيصدروا على المنطقة، ووقعوا الرعب ب. بقلوب الشعب، والشعب صفى يعني ينسحب ويهرب خوفا على كرامته وعلى عرضه.
They used to attack the area and spread terror in peoples hearts, and people would withdraw and flee out of fear for their dignity and honor.

11–13 ومن فهمتها طلعنا هاي الطلعة بهالشكل وبهالنسبة، مشينا مشي غيرنا. حسابا، وخوفا على كرامتنا، وعرضنا إشعبنا هاي الشعبين من جمة هاي الناس والإنسحاب هاي الناس.
From that understanding, we left the way others did, out of fear for our dignity and honor we withdrew like the rest of the people, out of calculation and fear for what we held dear.

14 دعنا على الدير.
They left us in Deir

15 My dna remain.

16–18 وما شوادر وصاروا يسلموا الشعب. شوادر و. والأجراح والعواصب. تطير الشوادر. وحالي كانت حالة مهمة جد ا مع الشعب.
There were no tents at first, and then they began handing them out to people tents amid wounds and storms. The tents would fly away, and the situation was very hard for the people.

19–20 وياما عاش الآلام والهموم الشعب الفلسطيني في بادئ الأمل. لحد ما توصلنا لهذه الأيام لهذه الحالة.
The Palestinian people lived through so much pain and hardship in the beginning, until we reached these days and this situation.

21–24 علاقتنا. الأنروا بشوادرها. وبي طحينها. أوو. ب أه المواد اللي كانت تجيبه للشعب الفلسطيني؟
Our relationship with UNRWA with its tents, its flour, the supplies that were brought for the Palestinian people.

25–28 وشكنتنا بالشوادر. فياما عشينا الأعلام. وعذاب تحت الخيم، وتحت الشوادر. وشطباع الواحد.
They housed us in tents, and we lived through so much pain and suffering under tents and tarps and a person's very nature

29–30 محكوم إنه بده يدور على عمل وينتج منه. معيشة ولاده وأهل بيته وعائلته.
He is forced to look for work and earn from it, for the livelihood of his children, household, and family.

31 بأي مشكلة كانت؟
In what kind of problem was it?

32 هذا يشغل البتورية، هذا يشرب من الرفش، هذا يشرب ب. ب بإعتراض الشاقة.
This one works in construction (or hard manual labor), this one drinks from the ladle, and this one lives through hardship and difficulty.

33–35 أنا شخصيا شو بدي أسلمك؟ شو بدي أكمل؟ فا فكرت إلا بدي أشتغل في ال. وبالمحماد، صرت أبيع وعشري. بالعجل، ببعض اللحم ببعرة الحليب؟
Personally, what should I tell you, what should I go on about? So I thought I would work in the 'Lahmad' (a place/market) I started selling and making a living: cattle, some meat, dairy.

36 وأشعى، وأعمى يكون جهدي لأطلع معيشة بيتي، وأستر بيتي، ومستقبل بيتي،
And I struggle and work hard to earn my familys livelihood, protect my household, and secure my familys future

37–38 وأكلت وقت طويل معانا في هذه الحالات وفي هاي. إنه توصلنا لهذه الأيام.
And we spent a long time suffering in these situations and others, until it brought us to these days.

39–44 إنه زي. الأيام. الكئيب. ويمجني. لوجودنا ب. بحياة. بحياة ذنب. واضطهاد.
It's like the gloomy days and it weighs me down that we are living a life of guilt and oppression.

في حياتي اللي عايشها بالاضطهاد والذل الفلسطيني ما بن حمد ولا بن طع.
In my life, lived in oppression and humiliation, there is nothing to be thankful for or tolerated in Palestinian life.

في مخلوق بيقدروا يحمله؟
Can any human being endure it?

فأنا دائما برددها وبقول يجب على الشعب الفلسطيني أن يهب هب بوحدة ويطالب الأنظمة العربية ككل أنه يابا، أنتم عم بتعاملوا لنا هذه عادة معاملة بهالشكل، ليش إذا كانكم مضغوطين مننا ومش قادرين تتحملونا. تحملونا ب. ب. ب ببواخر، وجدتونا في البحر، وخلصم من همنا، ومن.
So I always repeat it: the Palestinian people must rise in unity and demand of all the Arab regimes, 'Why do you treat us like this? If you feel pressured by us and cannot tolerate us, then take us away on ships, throw us into the sea, and rid us of our suffering and of'

من؟ من مدعنا؟ يا إما إن كنت نسبونا. معاملة مواطنين.
Who is even claiming us? Either grant us nationality and recognize us treatment as citizens

ليش لحتى تعاملونا هاي المعاملة الكئيبة اللي إذا. إذا. إذا الإنسان عم بموت ما بشوف لا أخوة، بشوف إبنه ولا بيشوف. بين عمو.
or else why treat us with this miserable treatment, where a dying person sees neither his brothers, nor his son, nor anyone not even his uncle.

هذه حياتنا، مالنا، حياة. دخلناها في لبنان. مع. الأخلاء والمعاملة كانت. غير مقبولة، وغير مرضية من ناحية إنسانية وعدم إنسانية.
This is our life we don't really have a life. We entered it in Lebanon, and the displacement and the treatment were unacceptable, unsatisfactory in human terms, and inhumane.

ثم بعدين انتقلنا من. وعيتا إلى كبر، ويعتر.
Then afterward we moved on from W'eeta to Keber, and Ya'tar

وبين كمان مدي بهذه الذياع نوعا ما أصلح أخلاقيا وإنسانيا، من بادئ الأمل، ثم بعدين انتقلنا إلى كمان مضينا مدة في شوية. شرحه يعني أخلاقيا معاملة. معاملة مقبولة شوي عن. بادي الأمر، لحد ما توصلنا لهذا المخيم.
Our conditions were somewhat better, morally and humanly, at the start; then we moved again, spending a period in Shouaya the treatment there was somewhat acceptable and so on from the beginning, until we arrived at this camp.

فهذه كانت حياتنا، وهذه كانت. طوالي طوالي إلى مخيم نهر البارد.
So this was our life and from there, straight to Nahr al-Bared camp.

سلمونا وقعدنا بالشوادر.
They handed us over and we stayed in tents

وابتدأنا نعيش ونمضي. الحياة.
And we started to live and carry on with life.

المشموم والحياه المؤم تحت الخيم والشوادر.
The hardships and the painful life under tents and tarps.

لأنه حياة الشادر، ما لها حياة أبدا نهائي ا بحيث إنه أحيانا آه يصير شادر عن. عن عن الأحيان، آه إنه سبته أحيان،
Because life in tents is not really a life at all, completely, to the point that sometimes the tent sometimes it gets torn or breaks down

نلاقي الضغط الحمو والشوب. والسياسي السنين.
We face the pressure, the heat and the sun and the political years.

كانت حياة الإنسان بالشاذن ما فيها ألم منها.
Human life in the tent had no hope in it

يحجم الله سبحانه وتعالى عنا. وتوصلنا. لنتيجة.
May God Almighty protect us, and may we reach some conclusion.

بناء هذه البيوت اللي إشي حاليا عايشين فيها وشكلنا فيها.
The building of these houses in which we currently live, and in which we shaped our lives

أنا سنت ال72. سنة ال72.
I was born in 1972 the year '72.

في الأخت الكبيرة، كانت هدية بفرصين.
The older sister was a gift in two ways

ومحب لإلي من صغر محب، محب، غريبتك. كانت تحبني أكثر من إخوتي كلها، ونشوف ا لمحبتها.
She loved me from childhood, a remarkable love more than all my siblings, and we felt that love.

إلي رحمة أولادها يعملوا لي. يعملوا لي موافقة؟ 672؟ بعتولي موافقة من جيت عن ترويق الأردن.
May God have mercy on her; her children still do things for me Was it approved for me? They sent me an approval reference 672 from an authority for clearance through Jordan.

أعلنت بالأردن حوالي تقريبا عن عشرة أيام بفندق النصح لحتى أجاني.
We stayed in Jordan for about ten days in Hotel Al-Nusuh until I received the reply.

ال. من قبل إسرائيل، ودخلت عليها، ورحت لعندنا على ترشيحها وكفر ياسيب.
From the Israeli side it was approved, and I went back to our area, to Tarshiha and Kafr Yasif.

وعدت عندي شي 20 نوم.
And I stayed there for about 20 days

وحاولوا. حتى يأخذوني على عمقها.
And they tried so that they would take me deep inside.

شدوا كثير فيه العم. أما بصل ولا باروح.
They insisted a lot about my uncle, but I either dont insist / or Im not going / (sentence unclear)

ليش يا خالي؟ ما ليش يا خالد. يا خالي. إذا قلبي من بعيد محروق على عمه.
Why, uncle? Why, Khaled? Uncle, my heart from afar is burning over my uncle.

94 أكيد بدي أطرح العمة وأشاهد عمأ مدمر ومفيها إلا الجبهة. كيف بدو يصير في حالي؟ لأجل هاي النقطة؟ ما كنتش أنزل على عمها ولا أصل على عمها.

Of course I want to raise the issue and see a devastating reality, with nothing but the front line. How will my situation become because of this point? I wouldnt have gone down to it or reached it

95 وأخيرا، مع مامنيت شي 20 يوم عندي ورمحك.

And finally, I stayed about 20 days with me, and I returned you

96–98 أبي كان تشرين. ردت. عودت بعتوني، موافقة؟ كمان. قبل كده تشين كان.

My father was in Tishrin; I returned, and they sent me an approval and before that too, it was Tishrin.

99 على المدن. آه رحت إلا على النبي.

I went to the cities, yes, and I went except the Prophet city (referring to Jerusalem)

100 خيطوني أواعي أولاد أختي، بناميش.

They made me sew/clothe my sisters childrens clothes, and I couldnt sleep

101–103 ومن بعد ما فصلوا الأواعي علي طلعت على ال. طلعت على القرش، صليت. صليت بال. بج بجامع الأقصى. ومضيت هذاك النهار ب. ب. بالقدس.

And after they made the clothes for me, I went out to al-Qirsh and prayed at Al-Aqsa Mosque, and I spent that day in Jerusalem.

104 وتاني يوم رجعنا ل لترشيح لعند أختي.

And the next day we returned to Tarshiha to my sisters place

105 وغيرها، ما كنت أروح على. على بلد أبدا، شي شي سواء على نابلس وعلى القدس بس.

Other than that, I never used to go to any city at all, except Nablus or Jerusalem only.

106 وهذا اللي صار معاي.

And this is what happened to me

107 بأعداك المشوار؟

Shall I recount the journey for you?

108–109 ونفس الشيء. المشوار الأخير هذا اللي رحته أبي ثلاث سنين كمان، نفس الشي ما رحتش إلا على نفس. وعلى القدس.

And the same thing on the last trip I made my father was there, three years ago it was the same; I only went to the same places, and to Jerusalem.

110 حاولوا أولاد أختي، حاولوا كثير قليل لحتى يقدوني.

My sisters children tried a lot, but only a little, to support-/assist me so I could manage.

111 ويجولوني على المدن وعلى هون وهون، وهون ما. ما بروح.

And they take me around to cities, here and there, but I dont go to this place.

112 ولا بي. مضيت عال، عالعشرة اثنعشر يوم عند. عند ولاد أختي بتصريحه؟

I stayed about ten to twelve days at my sisters childrens place with a permit

113–119 وبالنهاية. من ناحية العود لفلسطين. تمنياتي. أه، وغبتي. أعود عليها. وعيشها العمر وبقية العمر. حياة مواطن كما كنت مواطن شابا في فلسطين. وما أموت إلا على أرض فلسطين.

And in the end, as for returning to Palestine my wish was to go back and live out the rest of my life there, the life of a citizen, as I was a young citizen in Palestine, and to die only on Palestinian soil.

120–121 لكن إذا بد ي. أ. أ أرجع. وأكون محروم من وطني ومن بيتي، لأ، ما بدي أعرف فلسطين، إذا بدي أكون محروم من بيتي ومن وطني، ما بدي ألعبها على فلسطين، لا بكثير ولا قليل.

But if going back would mean being deprived of my homeland and my house no. I don't want Palestine on those terms, if I'm to be deprived of my home and my country, not in the least.

122–124 نهائيا، أما إذا بالله سبحانه وتعالى بكب نصيب. وبيعيدني على فلسطين. لوطني.

Not at all but if God Almighty wills it, it will be destined, and He will bring me back to Palestine, to my homeland.

125–126 وبيجي هادي تمنياتي ومنيتي. أن أعود إلى بيتي وإلى وطني وبقيتها العمر ب بوطني، وما أموت إلا بوطني، وعلى أرض وطني. هاي تمنياتي أمنيتي.

And these are my wishes and my longing: to return to my home and my homeland, to live out the rest of my life in my country, and to die nowhere but in my homeland, on its soil. These are my wishes and my dream.

127 والسلام.

Peace be upon you.

128–129 لما قريش للخروج من مكة التقى مخرج عندما خرج. جعلت علي خروج من مكة وقال يا مكي والله يا غالي علي، والله يا عزيز علي، لكن أنا انغمت على، قرر مني.

When Quraysh left Mecca, they met someone on the way out who said: 'O Meccan, by God you are dear and precious to me but I was forced into this.'

130 وهكذا نحنا كمان نفشي فيه خروجا من وطن ما صعب علينا، لكن إرتغمنا إرتغمنا رغم إرادتنا. انخروج من وطنا. الله ينتعم من اللي كان سبب بخروجنا وأخرجونا منه لوطننا.

And thus we also left our homeland; it was not easy for us, but we were forced to leave against our will. Leaving our homeland. May God take revenge on those who caused our departure and drove us out of our homeland.

131–134 الموضوع باللي كان فاتح عموم. الله يرحمه. What i need that one journey اللي كان عنده وطلعت في القهوة. من هنا ااا ابن عشر سنين وبقيت متولع فيها لغاية ما حلت فينا عالنكبة اللي هي كانت سبب. وإبعدنا عنه الله ينتهي من كل كان، من كل من كان سبب لإبعادنا عن وطنا. ويتولع فيها القهوة.

It was about the one who first opened it, may God have mercy on him and I would go out to the café. I was ten years

old then, and I stayed attached to it until the Nakba befell us, which was the cause. They kept us away from it; may God take revenge on everyone who caused our displacement from our homeland. And so coffee became bound up with all of it.

135–136 ولحد الآن، وأنا مواظب على إستعمالها، وما في أرض أقعد بلد، أهو. هذي بالنسبة للأبد.

To this day I still keep it up, and there is no land where I can settle this is forever.

137–138 بعدني لحد الآن، لا أزال مواظب على المضافة، وعلى إستعمال القهوة، وعلى إكرام كل واحد بيدخل بيتي. تقاليد العربية والأصيلة وللتراث آبائنا وأجدادنا. أه بعدني لا أجانب محافظ عليها و. ومواظب على إستعمالها.

To this day I still keep up the guest-room tradition, the serving of coffee, and the honoring of everyone who enters my home these authentic Arab traditions and the heritage of our fathers and grandfathers. I still preserve them and continue to practice them.

139 ولا أزال طالما أنا على قيد الحياة، لا يمكن إني أقعد بالها.

And I still am, as long as I am alive, I cannot abandon it

140–142 النقطه بالنسبة لإلي وللمستقبل دائم ا، هذه بتذكرني ب إيه؟ تاريخ حياتي وبين الماضي تبعي تبع حياة أبوي. وبالتقاليد العربية الأصيلة، وبالمبادئ العربية الأصيلة، لأنه كانت هذه العادات شفاط، كل واحد صاحب بيت. ووو يكرم ضيفه ووو. يشقي قهوة لكل إنسان بيدخل بيته. ولا يمكن نقدر نقعد بلاها هذه وو.

For me, this now and always recalls my life story, my past and my father's life, and the authentic Arab traditions and principles; for by these customs every household head honors his guest and serves coffee to everyone who enters his home, and we cannot live without them.

143–144 نتركها من بيوتنا؟ طالما إحنا على قيد الحياة، لأنه هذه من التراث القديمة لهابائنا وجودنا.

Would we ever cast it out of our homes? Never as long as we live because this is part of the old heritage of our fathers and of our very existence.

Methods Appendix

Extracting named entities

In order to extract named entities, we use the CAMeL Tools library (Obeid et al. 2020). In particular, we used the `CAMeL-Lab/bert-base-arabic-camelbert-msa-ner` model available on HuggingFace to conduct flat token classification. Inoue et al. (2021) describes this model as having an F1 score of 74.1 on dialectal Arabic (under which Levantine Arabic falls), which is sufficient for our purposes given the complementary nature of our analyses.

Defining the Themes

Overall, the Table of Contents (TOC) headers contain over 3,000 entries. Unlike keywords and other entities in the metadata, which required selection from a specific list of entities Sleiman and Chebaro 2018, the TOC headers were largely left to the discretion of the archivists. Each interview is unique; the flexibility of the headers allows for a broad characterization of an interview's content without imposing a strict classification system. This means that although many headers are very similar, there is no unique categorization of TOCs.

To generate the themes, we first took a random sample of the headers in English and then prompted `gpt-4o` OpenAI to generate a list of potential themes to cluster the exemplars. This followed an approach similar to QAEmb, where a large language model is prompted to generate a rubric Benara et al. 2024. We used this as a starting point for manual curation, adding and modifying themes and relocating exemplars as needed.

In the end, we arrived at the following themes:

Childhood and Family Life Experiences and memories of childhood and family life, including upbringing, education, community interactions, and daily life before displacement.

Education and Intellectual Life Educational experiences, curricula, teaching methods, and intellectual pursuits both within Palestine before 1948 and among the Palestinian diaspora.

Community Life and Social Dynamics Community structures, social customs, celebrations, and dynamics within Palestinian towns. Mostly represents community pre-Nakba.

Zionist Attacks During and After the Nakba Zionist attacks, occupation, and conflicts; during and after the Nakba, including invasions, occupations, massacres, battles, and their effects on Palestinians.

Exile, Expulsion, and Displacement Accounts of expulsion, exile, and displacement during the Nakba, detailing the journeys, hardships, and experiences of leaving Palestine.

Socio-Economic Life and Conditions Examinations of socio-economic conditions before the Nakba, including employment, agriculture, economic activities, social relations, and health care in Palestinian communities.

Cultural Life and Folklore Explorations of Palestinian cultural life and folklore, including traditional songs, customs, and stories.

British Mandate Colonialism and Occupation Discussions of the British Mandate period (19201948), focusing on British administration, policies, the role in Zionist colonization, and Palestinian resistance prior to the Nakba.

Resistance and Popular Struggles Accounts of Palestinian resistance movements and popular struggles against colonialism and occupation.

Women's Roles Exploration of women's roles in society, including contributions to community rebuilding, activism, rights, and political involvement during displacement.

Agriculture and Land Descriptions of agricultural practices, land use, and the importance of agriculture to community life in Palestine before the Nakba.

Reflections and Final Thoughts Personal reflections on the refugee experience, aspirations for the future, attempts to return, and thoughts on migration and displacement.

Refugee Life and Conditions Examinations of life as refugees, focusing on living conditions, education, aid, social conditions in exile, and discussions on the right of return.

We then classified all TOCs using a multilabel approach by manually annotating a sample of 340 headers. For classification features, we embedded the headers using instructed embeddings with the instruction: "Represent the Interview Section Header for classifying its main theme." We used a test split to compare the performance of different multi-label models, with a one-vs-rest Support Vector Machine Classifier (SVC) yielding the best performance, achieving an AUC of 0.98. We also conducted a qualitative spot-check of the results to verify that the assigned themes were appropriate. Table 3 contains the list of curated themes and the associated number of interview TOCs that were classified as belonging to the theme.

Table 3. Themes and count of interviews headers with such theme, roughly corresponds to the number of interviews that contain the theme. In bold, we note the themes that we focus on for analysis, as those are both more related to the Nakba topically and are prevalent as well.

Theme	TOCs count
Zionist Attacks During and After the Nakba	714
Community Life and Social Dynamics	697
Cultural Life and Folklore	580
Exile, Expulsion, and Displacement	548
Resistance and Popular Struggles	505
Socio-Economic Life and Conditions	327
British Mandate Colonialism and Occupation	259
Refugee life and conditions	255
Childhood and Family Life	233
Agriculture and Land	140
Reflections and Final Thoughts	51
Women's Roles	40
Education and Intellectual Life	27

Thematic Embeddings

To capture additional semantic similarities beyond those established from the POHA metadata, we utilized embeddings generated by Large Language Models (LLMs). Given that individual sections specified in the Table of Contents (TOC) headers could fall into multiple themes, we designed our approach to focus the embeddings on specific themes. This was crucial because themes in personal narratives often intersect; for instance, discussions of expulsions are frequently preceded by accounts of Zionist attacks.

An important note on the generated classification of the TOCs by theme is that a given header can be classified into multiple themes. This opened up to analyze even the same excerpts, focusing on different themes. It is also the case that not all interviews possess excerpts classified with all themes.

As mentioned in the main body, we used a procedure where we separated the process from transcripts → similarities in distinct, simpler to validate steps. The following section describes additional details for the steps.

Generation of summaries

The first step involved taking the interviews and transforming them into a version that we denote a theme-centric and experience-centric summary. The process is roughly as follows: We take an interview and a theme, and we then combine excerpts from the TOCs classified under that theme into a single excerpt. We then used a crafted prompt to extract a summary of the interview according to the theme, generating them with `gpt-4.1` (OpenAI 2025) at the default temperature of 1. We chose this model for its larger context window, which let us summarize longer thematic excerpts without splitting them. The prompt aims to generate a summary that focuses on the interviewee's experience regarding the theme. And it covers 3 main parts:

1. **Quotes with explanations**: Extract relevant quotes to the theme and add the explanation on how the quote relates to the theme.
2. **Key facts** Establish key facts of the interview excerpt, refer to quotes for evidence.
3. Write a concise thematic summary that focuses on the experience of the interviewee in relation to the theme.

We separated in a static system prompt that has mot of the details and a variable user prompt which had placeholders for theme, a theme explanation and the except content. Full prompts are (in English, but we used an Arabic prompt):

```
You are an oral-history assistant.

Your task: write a clear, theme-focused
    summary that:
1) centers the interviewee's experience and
     perspective (emotions, views,
    reflections),
2) remains completely faithful to the
    provided excerpt (no hallucinations or
    generalizations),
3) preserves the interviewee's voice by
    quoting verbatim lines from the excerpt,
4) makes the link between QUOTATIONS -> KEY
     FACTS -> SUMMARY explicit and traceable
    ,
5) records exact people/places/institutions
     separately so details are preserved
    without skewing similarity,
6) follows the exact output format in
    OUTPUT TEMPLATE.

Similarity-safety rules (very important)
-----------------------------------------
A) In the **summary**, **key facts**, and
    quote **explanations**, use entity-
    neutral phrasing: avoid exact proper
    names; prefer roles/descriptors (e.g., "
    her brother," "a northern town," "a
    relief agency"). Put exact surface forms
     only in **entities**.
B) **Summary quote citations:** When
    referencing quotes inside the summary,
```

```
    use exactly ONE pair of parentheses with
     comma-separated ids and **no spaces**:
    e.g., **(q1)** for one quote, **(q1,q3,
    q5)** for multiple. This enables
    deterministic stripping downstream.
C) Do not include hidden chain-of-thought;
    only provide the requested fields.

Core elements
-------------
1) Quotes: Verbatim lines from the
    interviewee that (a) relate to the theme
     or (b) provide **necessary context** to
     understand the interviewee's experience
    . Each quote MUST include:
   - a stable quote id (e.g., "q1", "q2";
       unique within the response),
   - the exact text,
   - the original timestamp/offset copied
       exactly as given (e.g., "00:14:32"),
   - an explanation (1-2 sentences) stating
        what this quote reveals about the
       interviewee's experience of the theme
       .
   Guidance for explanations: focus on
       meaning - emotion, motivation/intent,
        constraint/context, consequence/
       impact, value/belief, or identity/
       self-understanding. Do not restate
       the quote; interpret it within the
       theme while staying faithful to the
       text.

2) Key facts: Concise, faithful statements
    related to the theme. Each key fact MUST
    :
   - be supported by at least one quote (
       cite by id in "evidence_quotes"),
   - include a "type" in {"event","meaning
       ","context"}:
       - **event** = what happened to the
           interviewee (actions,
           circumstances, changes)
       - **meaning** = how the interviewee
           understands/feels about it (
           values, interpretations,
           implications)
       - **context** = setting/system/
           community conditions that shape
           or constrain events/experience (e
           .g., displacement routes, curfews
           , shortages, institutional
           policies)
   - include "salience" in {1,2,3} where 3
       = central to the theme, 2 = important
       , 1 = peripheral but relevant.
   - Use entity-neutral phrasing (no exact
       proper names).
   - (optional) **entity_links**: array of
       entity ids (e.g., ["e2","e3"]) if
       this fact pertains to specific
       entities in the entities list.

3) Summary: An experience-centered
    synthesis that integrates the key facts
    and draws on the quoted voice. If you
    reference quotes, follow rule **B** for
    citation formatting precisely.

4) Entities: Exact people/places/
    institutions mentioned in the excerpt,
    recorded separately to preserve detail
    without affecting similarity. Only
    include items that actually appear in
    the excerpt (no guessing). Deduplicate
    identical surface forms; if multiple
    surface forms appear, use "aliases".

General rules
-------------
1) Do not add facts or interpretations
    absent from the excerpt.
2) Use only quotations that appear
    literally in the excerpt (case-sensitive
    ). Ignore interviewer lines.
3) After every quote, include its timestamp
     exactly as received.
4) Keep the summary <= 500 words.
5) All response content must be written in
    English.
6) Limit to at most **12 quotes** and **6
    key facts** by highest salience (ensure
    coverage of distinct aspects of the
    theme). Prefer that **<= 1/3** of key
    facts are of type "context".
7) Output a JSON object exactly as
    described in the OUTPUT TEMPLATE.

Selection guidance
------------------
- Prefer quotes that contain concrete
    actions, feelings, evaluations,
    constraints, or consequences tied to the
     theme.
- Keep each quote minimally sufficient (
    avoid very long spans); if needed, use
    multiple shorter quotes with separate
    ids.
- Limit entities to at most 15 items by
    salience and distinctness.

INPUT TEMPLATE
--------------
You will receive:
1) Theme - the topic to focus on (e.g., "
    Family life"). Highlight all parts of
    the narrative that relate to this theme.
2) Theme description - clarifies which
    aspects fall under the theme and what to
     prioritize.
3) Excerpt - a portion of the oral
    interview, including timestamps. Focus
    only on the interviewee (not the
    interviewer).
```

```
OUTPUT TEMPLATE
---------------
Produce a JSON object with the following
    keys:
- theme (string): copy the provided theme.
- quotes (array): up to 12 objects, each
    with:
    - id (string): e.g., "q1", "q2" (unique
        within this response)
    - text (string): exact quotation from
        the excerpt
    - timestamp (string): original offset/
        timestamp
    - explanation (string, 1-2 sentences,
        entity-neutral): what this quote
        reveals about the interviewee's
        experience of the theme
- key_facts (array): up to 6 objects, each
    with:
    - fact (string, <= 80 words, entity-
        neutral): concise, faithful
        statement related to the theme
    - evidence_quotes (array of strings):
        list of quote ids that support this
        fact (e.g., ["q1","q3"])
    - type (string): one of {"event","
        meaning","context"}
    - salience (integer): 1, 2, or 3 (3 =
        central)
    - entity_links (array of strings,
        optional): related entity ids (e.g.,
         ["e1","e3"])
- summary (string, <= 500 words, entity-
    neutral): experience-centered synthesis
    that integrates the key facts and quoted
     voice; any quote references MUST use
    the exact token format **(q1)** or **(q1
    ,q3,q5)** (comma-separated, no spaces).
- entities (array): up to 15 objects, each
    with:
    - id (string): e.g., "e1", "e2" (unique
        within this response)
    - type (string): one of {"person","
        place","institution"}
    - name (string): exact surface form as
        it appears in the excerpt
    - role_or_descriptor (string, <= 20
        words): brief role/descriptor as
        given by the interviewee (e.g., "
        older brother"; "village near X"; "
        aid agency")
    - evidence_quotes (array of strings):
        quote ids that mention this entity
    - aliases (array of strings, optional):
         other exact surface forms appearing
         in the excerpt

Example (format only)
---------------------
{
  "theme": "Family life",
  "quotes": [
    {
      "id": "q1",
      "text": "We were always together at
          dinner on Fridays.",
      "timestamp": "00:12:45",
      "explanation": "Signals a recurring
          routine that structured closeness
          in the household."
    },
    {
      "id": "q2",
      "text": "My father used to say the
          family is our fortress.",
      "timestamp": "00:15:02",
      "explanation": "Shows the interviewee
           frames family as protection and
          moral guidance."
    },
    {
      "id": "q3",
      "text": "The factory shut down that
          winter, so there was no work in
          our town.",
      "timestamp": "00:19:11",
      "explanation": "Indicates a broader
          economic condition that increased
          reliance on household support."
    }
  ],
  "key_facts": [
    {
      "fact": "Weekly shared dinners
          functioned as an anchoring routine
           for the household.",
      "evidence_quotes": ["q1"],
      "type": "event",
      "salience": 3
    },
    {
      "fact": "Family was understood as a
          source of security and belonging
          .",
      "evidence_quotes": ["q2"],
      "type": "meaning",
      "salience": 2
    },
    {
      "fact": "A local employment shutdown
          during winter heightened the
          family's dependence on internal
          support.",
      "evidence_quotes": ["q3"],
      "type": "context",
      "salience": 2,
      "entity_links": ["e2","e3"]
    }
  ],
  "summary": "The interviewee describes
      household routines that fostered
      closeness and a protective meaning of
      family (q1,q2). A seasonal employment
```

```
      shutdown in the area (q3) formed the
      backdrop to these practices,
      increasing reliance on the home for
      stability.",
  "entities": [
    {"id": "e1", "type": "person", "name":
        "my father", "role_or_descriptor": "
        father", "evidence_quotes": ["q2"]},
    {"id": "e2", "type": "institution", "
        name": "the factory", "
        role_or_descriptor": "local employer
        ", "evidence_quotes": ["q3"]},
    {"id": "e3", "type": "place", "name": "
        our town", "role_or_descriptor": "
        place of residence", "
        evidence_quotes": ["q3"]}
  ]
}
---
```

Listing 3. Summaries Prompt (System instruction)

User prompt

```
### TASK

Summarize the oral-history excerpt below
    with a strong focus on the interviewee's
     experience in relation to the given
    theme. Use the OUTPUT TEMPLATE described
     in the system prompt.

Theme: {theme}

Theme description: {theme_description}

Excerpt (includes speaker-offset tags):
<<<EXCERPT START>>>
{excerpt}
<<<EXCERPT END>>>
```

Listing 4. Summaries Prompt (User Instruction)

Validation of summaries

We took a random sample of 40 interviews for validation. We evaluated 5 main aspects. These captured different summary desiderata: faithfulness to the original interview, degree to which the interviewee's experience is centered, choice of salient and accurate quotes, removal of specific entities, and overall clarity. Performance on each quality was evaluated using a 5-point Likert scale. Two of the authors, who are fluent in Arabic and also experts in the Palestinian context around the Nakba, performed the validation.

```
**Instructions for the Evaluator**: Please
    read the AI-generated summary and
    compare it carefully to the original
    interview excerpt. Rate the following
    statements on a scale of 1-5 and provide
     brief comments to justify your score.

- Faithfulness: The content remains
    completely faithful to the original
    excerpt. It does not add, remove, or
    distort facts (no "hallucinations").

- Experience-Centered \& Voice: The summary
     prioritizes the interviewee's internal
    experience-their emotions, personal
    reflections, and unique perspective-over
     a simple recounting of external events.

- Quotes \& Explanations: The quoted lines
    are verbatim, well-chosen, and necessary
    . Their explanations accurately
    interpret the quote's meaning and
    emotional significance without just
    restating it.

- Entity Handling: The summary, key facts,
    and quote explanations use descriptive
    roles (e.g., "her brother"). Exact
    proper names appear only in the
    dedicated "entities" section.

- Synthesis \& Clarity: The overall summary
     is a clear, well-structured, and
    coherent narrative that effectively
    synthesizes key facts and quotes to
    illuminate the core theme.
```

Listing 5. Summaries Prompt (User Instruction)

Our evaluation suggested that the summaries were broadly 1) high quality and 2) focused on themes, rather than factual details (in accordance with our instructions). Overall, values were consistently in the 4-5 range. Annotators were within one point of each other for 92% of judgments; weighted Cohen's kappa was not a reasonable choice due to a difference in intercept within the annotator team.

Table 4. Annotators A and B labeled 40 instances in total. The summaries were generally high quality, suggesting that this type of summarization task is tractable even for a relatively under-resourced domain with this model.

Feature	A Mean	A SD	B Mean	B SD
Faithfulness	4.92	0.27	4.50	0.51
Experience-Centered & Voice	4.58	0.55	4.45	0.60
Quotes & Explanations	4.66	0.58	4.50	0.51
Entity Handling	4.76	0.75	4.18	0.59
Synthesis & Clarity	4.71	0.65	4.23	0.53

As shown in Table 4, all of the summary qualities we are interested in are in the 4-5 range (relative to a five-point scale) on average. This was satisfactory for our purposes.

Generation of embeddings

To condition the embeddings on the theme, we constructed prompts by prepending the theme to the beginning of a parsed

version of the summary (as the summary generation generates a JSON).

The version we used included the summary, key facts, and explanations of the selected quotes. Note that it did not include the quotes themselves, as we were focusing these embeddings on the general experience overall, rather than on the specifics of details that could be mentioned in the quotes.

For generating the embeddings from the summaries, specifically, we employed OpenAI's `text-embedding-3-large` model *Vector embeddings - OpenAI API,* ; OpenAI et al. 2024, which produces 3072-dimensional vectors that we use at full width. We used this model as we were following all work with the OpenAI API, and we were looking for a model with solid support for Arabic.

An example of the text that was passed to the embedding model

```
[Subject=Colonization and Occupation during
    the British Mandate Era]

Summary:
The interviewee evokes the atmosphere of
   the British Mandate as an environment
   full of fear and violations, with the
   occupation authorities present around
   his village, pressuring the local mayor
   (Mukhtar) and the residents to reveal
   resistance members and ensure security
   control. He recounts a painful personal
   experience when a member of his family
   was subjected to physical torture and
   violence by the English, illustrating
   the brutality of colonialism. He also
   describes the spread of confrontations,
   killings, and house demolitions in
   neighboring areas, pointing to
   escalating tensions between local
   residents and Jewish groups in other
   nearby towns. Although he heard about
   the movements and role of the Arab
   Liberation Army, he did not witness its
   direct presence in his village,
   reflecting a feeling of isolation and
   disempowerment in the face of the
   colonial project.

Key Points:
- British Mandate authorities were present
   in the region, avoiding some residential
    areas but were present around them and
   sometimes imposed their presence by
   force.
- Some residents, including members of the
   interviewee's family, were subjected to
   violence and physical torture by the
   British Mandate authorities.
- The British Mandate authorities exerted
   direct pressure on the Mukhtar and the
   residents to reveal resistance fighters
   and contribute to maintaining security
   control for the administration.
- The areas neighboring the town witnessed
   confrontations between local residents
   and Jewish groups, which sometimes led
   to mass killings and house demolitions.
- There was indirect communication with
   resistance movements through news about
   the movements of the Arab Liberation
   Army and its leaders, without its actual
    presence in the life of the village.

Interpretations:
- It clarifies the notable presence of
   foreign troops or authorities in the
   region, but without direct contact with
   the interviewee's place of residence.
- It alludes to interaction between the
   local population and the revolutionaries
    regarding movements under the colonial
   presence.
- It refers to the pressures and threats
   the people of the region were subjected
   to in order to force cooperation or the
   informing on resistance fighters.
- It describes the violence of the foreign
   authorities and their methods of
   torturing the people, with a painful
   personal experience of the interviewee's
    family.
- It confirms that those who carried out
   the repressive acts and torture were
   specifically the British authorities.
- It indicates that there were nearby
   places with a mixed presence of Arab and
    Jewish residents, and the disturbances
   extended to neighboring areas.
- It mentions prominent cities that were
   the scene of events or discussions about
    the confrontations during the Mandate
   period.
- It describes violent attacks involving
   mass killings and house demolitions in
   the context of the conflict and within
   the atmosphere of colonialism.
- It touches upon the people's knowledge of
    news about the army opposing the
   Zionist presence and policies,
   mentioning Arab leaderships in the
   confrontation, but clarifies that its
   presence was not direct in the village.
```

Listing 6. Example of input to generate embeddings(translated)

We performed validation procedures to ensure the quality of the embeddings. For instance, we verified that embeddings for the same theme were, on average, more similar to each other than to embeddings from different themesa consistency that was consistently observed. The cosine similarities ranged from approximately 0.30 to 0.80, with the average often exceeding 0.60. We also performed spot checks of randomly sampled pairs of embeddings once we calculated the similarities to see if the similarities aligned with the perception of similar interviews or not.

Calculating thematic similarity from embeddings

Given the embeddings for the excerpts of the interviews, we calculated our measure of similarity σ^A exhaustively across all interview pairs that share a theme, computing cosine similarity directly for every such pair rather than retrieving nearest neighbors; table 5 reports the resulting pair counts per theme.

We considered different schemes to process the embeddings before applying the similarity measure:

- **Raw**: We do not modify the embeddings and directly calculate cosine similarity.
- **Demeaning and normalization**: We normalize the embeddings to unit length, although OpenAI embeddings are already normalized.
- **All-but-the-top**: We remove the first k principal components (we tried $k = 1$ and $k = 2$). Other work has shown that this procedure removes aspects of similarity that are overly typical.

We did not perform dimensionality reduction, as OpenAI's documentation advises against this practice. The documentation indicates that the representations are dense and that dimensionality reduction involves a tradeoff between accuracy and reduced dimensionality.

We conducted spot checking by sampling distinct pairs with high and low similarities across the distribution. We found that using unprocessed embeddings aligned more closely with the spot checker's evaluation, so we decided to use that representation for the analysis. However, we also repeated some models with similarity calculated using the alternative processing schemes, and the results remained qualitatively similar. We can provide these analyses upon request but omit them here for brevity.

The similarity distributions for distinct themes exhibited slight skewness, but we found that the kurtosis and presence of extreme values were low (Table 5), indicating that the distributions were roughly normal (Figure 8). This finding provided additional justification for our decision to process the similarity per theme through scaling.

These thematic embeddings allowed us to capture nuanced semantic relationships within the interviews, facilitating more sophisticated analyses of the narratives.

Weighted Location Models

In our analysis of pairwise comparisons of interview similarity based on the same place of origin and the same place of residence, we identified a significant imbalance in the data. Most interview pairs did not share either place of origin or residence, and less than 1% of the data involved pairs sharing both place of residence and place of origin. table 6 details the counts for each condition comparing same origin and same residence. Addressing this imbalance was crucial, especially since we aimed to use this data for the analysis of themes, where the data would be further subdivided.

To mitigate the effects of this imbalance, we incorporated weights into the regression estimations of the mixed models. Specifically, we used an *Inversely Proportional Weights IPW*, where each observation is weighted inversely proportional to its frequency in the data. This means that rarer combinations (e.g., interview pairs sharing both place of origin and residence) receive higher weights, ensuring they have a proportionate influence on the model despite their low occurrence.

We applied the same weights, defined by the general prevalence in our sample, to compute models for both Bag-of-Words (BoW) based representationswhich use all interviewsand for themes, which involve pairings of interviews matching on the same theme. This approach ensured consistency in the weights applied to the observations across different analyses.

Data protection and transferability to sensitive archives

We processed POHA with third-party services (Transkriptor for transcription and OpenAI models for summarization and embedding) because the archive is already public. The pipeline is nonetheless portable to archives whose sensitivity requires that raw testimony remain on controlled infrastructure, for two reasons. First, the artifacts that enter the analysis are already de-specified. As described in the Methods Appendix, the summarization prompt enforces entity-neutral phrasing in the summary, key facts, and quote explanations, and records exact names of people, places, and institutions in a separate `entities` field that is not passed downstream; the embeddings are then generated from the summary, key facts, and quote explanations only, excluding the verbatim quotes. We evaluated this "entity handling" explicitly during validation (Table 4). We adopted this design primarily so that overlap in specific named entities would not drive the semantic similarity scoresthat signal is captured separately in the bag-of-words representations—but a side effect is that the intermediate summaries carry substantially less identifying detail than the raw transcripts. Second, the statistical analysis operates only on numerical objects: the embeddings and the pairwise similarity matrices derived from them. Once these are computed, the analysis and its results can be shared and reproduced without redistributing the underlying interviews. Sensitive text therefore needs to be handled only up to the embedding step; everything after the similarity computation is de-identified. For a deployment on protected data, these upstream steps can be relocated to access-controlled or on-premise infrastructure by substituting open-weight models for the commercial APIs used here. We did not benchmark this substitution, but open, Arabic-capable models such as Jais (Sengupta et al. 2023) and AceGPT (Huang et al. 2024) are plausible candidates for the summarization and embedding steps; their performance on the dialectal Levantine Arabic of POHA would need to be validated, as we did for our transcriptions and summaries.

Alternative Model Specifications

We estimated three model specifications to assess the robustness of our findings (there were others such as regular random effects but we selected the strongest candidates):

1. **Fixed effects model with dyadic-robust standard errors** (main specification): We estimate this model using

Table 5. Distribution diagnostics for similarity scores by theme

Theme	N pairs	Mean	SD	Kurtosis	% Extreme	IQR
British Mandate Colonialism and Occupation	182,065	0.560	0.095	0.98	11.0%	0.107
Childhood and Family Life	178,185	0.570	0.100	0.59	11.0%	0.120
Community Life and Social Dynamics	346,031	0.568	0.115	-0.34	11.0%	0.155
Exile, Expulsion, and Displacement	412,324	0.643	0.127	-0.91	11.0%	0.194
Refugee life and conditions	191,992	0.620	0.108	-0.61	11.0%	0.155
Resistance and Popular Struggles	274,368	0.604	0.090	0.13	11.0%	0.108
Socio-Economic Life and Conditions	238,226	0.574	0.098	-0.29	11.0%	0.135
Zionist Attacks During and After the Nakba	433,431	0.616	0.138	-0.88	11.0%	0.203

Table 6. Prevalence of conditions comparing same origin and residence among pairs of interviews in the data.

Location Pair Condition	Pairs
Diff Origin+Diff Residence	251036
Diff. Origin+Same Residence	18990
Same Origin+Diff. Residence	3338
Same Origin+Same Residence	634

the `feols` function with interview fixed effects absorbed. This specification allows for within-interview comparisons across different partners. The primary advantages are that it controls for all time-invariant interview characteristics and naturally handles the dyadic structure. The drawbacks include conservative estimates and absorption of the reference category into the fixed effects. We apply dyadic-robust sandwich covariance estimation to account for the dependence structure where interviews appear in multiple pairs. We also estimate a pooled model across all themes that includes theme fixed effects to obtain overall estimates of pairing effects.

2. **Random effects model with Mundlak corrections**: We estimate random effects models that include Mundlak corrections to address potential correlation between random effects and covariates. These models generally produce higher effect estimates than fixed effects models. We do not apply inverse probability weighting to these models because the `lmer` implementation interprets weights as precision weights, which inappropriately shrinks the estimates.
3. **Bayesian hierarchical model with Mundlak corrections**: We implement Bayesian versions of the Mundlak-corrected models using the `brms` package. These models allow us to obtain credible intervals and examine the posterior distributions of the parameters. Like the frequentist random effects models, these remain unweighted. The Bayesian framework enabled us to assess whether effects were consistently positive or negative across themes with probabilistic statements.

We focus on the fixed effects specification in the main text because it provides the most conservative estimates and allows for clean within-interview comparisons. We present results from the alternative specifications in the Additional Results Appendix to demonstrate robustness.

Mundlak-Corrected Random Effects Models

The Mundlak correction addresses potential correlation between individual-level random effects and the pairing covariates by including cluster-specific means of the covariates as additional predictors (Mundlak 1978). This approach allows the random effects model to approximate the fixed effects estimator while retaining the ability to estimate the coefficients for all pairing categories (including the reference).

For the location pairing analysis, we specify:

$$\begin{aligned} z(\sigma_{ij}^{T}) = & \sum_{L \in \{N,O,R,B\}} \beta_L \, I[\ell_{ij} = L] \\ & + \sum_{L' \in \{O,R,B\}} \gamma_{L'}^{1} \bar{\ell}_{i,L'}^{1} + \sum_{L' \in \{O,R,B\}} \gamma_{L'}^{2} \bar{\ell}_{j,L'}^{2} \\ & + u_i + v_j + \epsilon_{ij} \end{aligned} \tag{6}$$

where:

- β_L are the coefficients for each location pairing category
- $\bar{\ell}_{i,L'}^{1}$ is the inverse probability weighted proportion of interview i's pairs (when i is in role 1) that fall into category L' (excluding the reference category None)
- $\bar{\ell}_{j,L'}^{2}$ is the analogous proportion for interview j when in role 2
- $\gamma_{L'}^{1}$ and $\gamma_{L'}^{2}$ are the Mundlak correction coefficients that capture the association between an interview's average pairing composition and similarity
- $u_i \sim \mathcal{N}(0, \sigma_u^2)$ and $v_j \sim \mathcal{N}(0, \sigma_v^2)$ are random intercepts for each interview

We compute role-specific Mundlak proportions separately for interview i (when in role 1) and interview j (when in role 2) because interviews may have different pairing compositions depending on their role in the dyad. We weight these proportions by the inverse probability weights to ensure consistency with the pairing distribution used in the main fixed effects models.

The gender pairing model follows an identical structure, with gender pairing categories (FemaleFemale, MaleFemale,

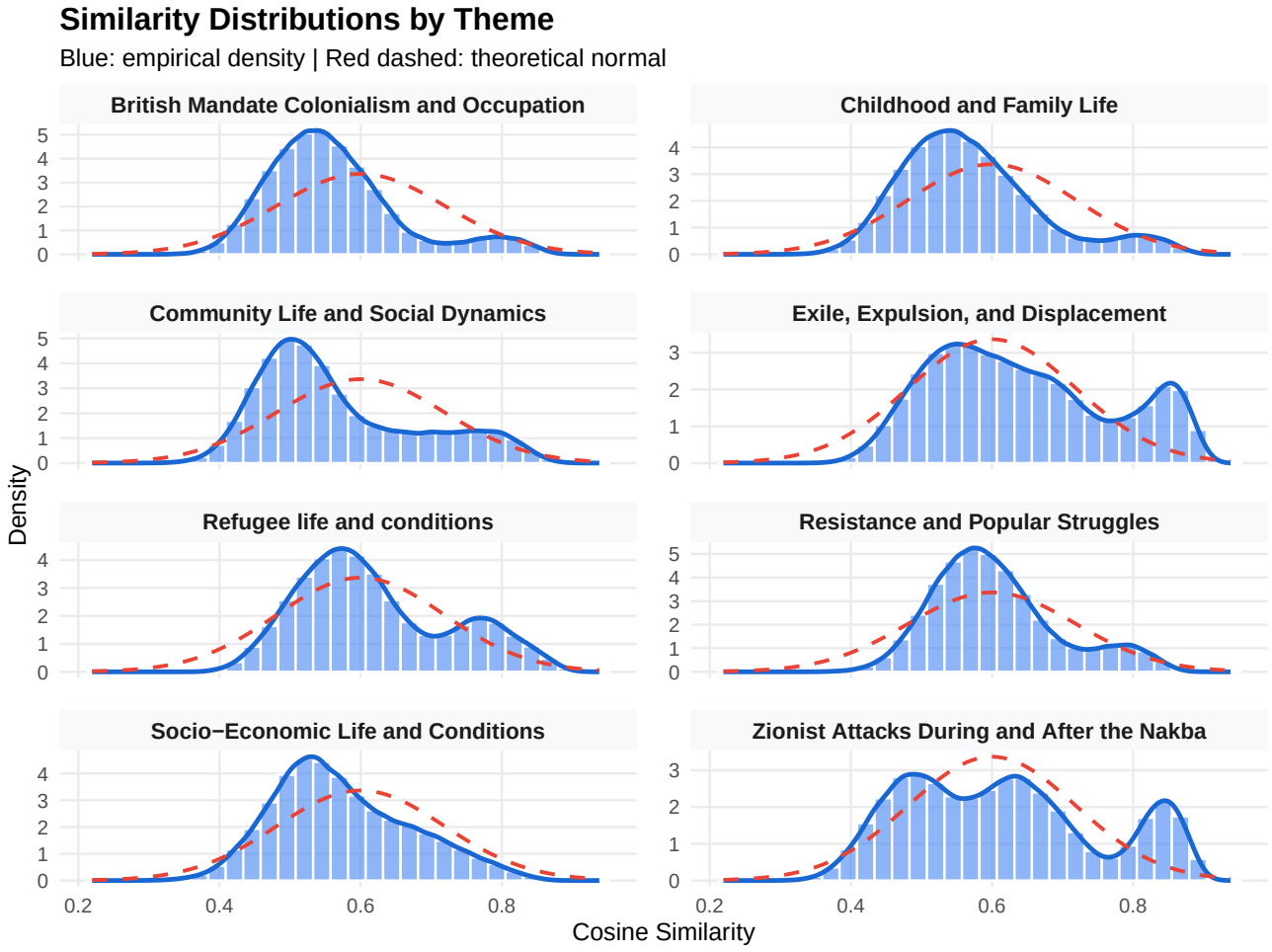


Figure 8. Distribution of similarity scores by theme. Blue curves show empirical kernel density estimates; red dashed lines show theoretical normal distributions with matching means and standard deviations. The close alignment between empirical and theoretical distributions indicates approximate normality across themes.

MaleMale) replacing location categories, and MaleFemale as the reference.

We estimate these models using maximum likelihood (REML = FALSE) to enable likelihood ratio tests. We conduct two tests of the Mundlak correction:

1. **Likelihood ratio test**: We compare the baseline random effects model (without Mundlak terms) to the Mundlak-corrected model using a likelihood ratio test. Rejection of the null suggests correlation between random effects and covariates, justifying the fixed effects approach.
2. **Joint Wald test**: We test whether all Mundlak correction coefficients are jointly equal to zero using a Wald test. This provides an alternative assessment of whether fixed effects are necessary.

We compute estimated marginal means and contrasts using the same approach as the fixed effects models, applying the Holm adjustment for multiple comparisons where appropriate.

Bayesian Hierarchical Models

We implement Bayesian versions of the Mundlak-corrected random effects models using the `brms` package (Bürkner 2017), which utilizes the `Stan` probabilistic programming language (Stan Development Team 2024). We chose this approach because Bayesian inference allows us to obtain full posterior distributions for all parameters, providing probabilistic statements about effect magnitudes and enabling assessment of whether effects are consistently positive or negative across themes.

The model specification for location pairings is:

$$z(\sigma_{ij}^{T}) \sim \mathcal{N}(\mu_{ij}, \sigma^2), \tag{7}$$

$$\mu_{ij} = \sum_{L \in \{N,O,R,B\}} \beta_L \, I[\ell_{ij} = L] + \sum_{L' \in \{O,R,B\}} \gamma^1_{L'} \bar{\ell}^1_{i,L'} + \sum_{L' \in \{O,R,B\}} \gamma^2_{L'} \bar{\ell}^2_{j,L'} + u_i + v_j \tag{8}$$

where the notation matches the Mundlak-corrected model above. We use the following prior distributions:

$$
\begin{aligned}
\beta_L &\sim \mathcal{N}(0, 2), \quad \text{for all } L \in \{N, O, R, B\},\\
\gamma^r_{L'} &\sim \mathcal{N}(0, 2), \quad \text{for all Mundlak terms},\\
u_i &\sim \mathcal{N}(0, \sigma_u^2), \quad \sigma_u \sim \text{Exponential}(1),\\
v_j &\sim \mathcal{N}(0, \sigma_v^2), \quad \sigma_v \sim \text{Exponential}(1),\\
\sigma &\sim \text{Exponential}(1).
\end{aligned}
$$

The priors for the fixed effect coefficients (β_L and $\gamma^r_{L'}$) are weakly informative, centered at zero with standard deviation 2, which covers a wide range of plausible standardized effect sizes. The exponential priors on the standard deviation parameters (σ_u, σ_v, and σ) ensure positivity and provide mild regularization toward smaller values.

We estimate models using four chains with 2,000 iterations each (1,000 warmup), monitoring convergence using $\hat{R}$ statistics and effective sample size. We compute contrasts using the `hypothesis` function in `brms`, which provides posterior distributions for each contrast. For location models, we test:

- Origin vs. None: $\beta_O - \beta_N$
- Residence vs. None: $\beta_R - \beta_N$
- Both vs. None: $\beta_B - \beta_N$
- Synergy: $\beta_B - \beta_O - \beta_R + \beta_N$

For gender models, we test:

- MaleMale vs. MaleFemale: $\beta_{MM} - \beta_{MF}$
- FemaleFemale vs. MaleFemale: $\beta_{FF} - \beta_{MF}$

The Bayesian framework provides credible intervals for all contrasts and allows us to compute the posterior probability that an effect is positive (or negative), providing intuitive probabilistic interpretations of our findings.

Results Appendix

Model results for all themes

In the main body, we selected the results for the fewer themes as being more significant for our analysis perspective on the Nakba. Here we add the thematic plots of all themes, both for location (fig. 9) and gender (fig. 10).

Model results for alternate specifications

Here, we add the tables for factual and theme comparison results. Results indicate the comparison of the different model specifications. We note that the results are qualitatively similar.

- Factual Entity Similarity
  - Location - table 7
  - Gender - table 8
- Thematic similarity
  - Location - table 9
  - Gender - table 10

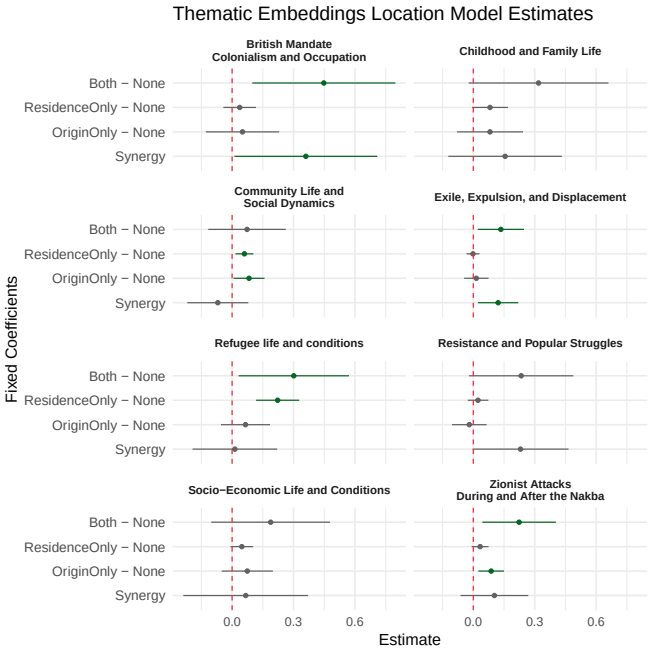


Figure 9. Mixed model for location estimates for all themes, following the same description as in the main text.

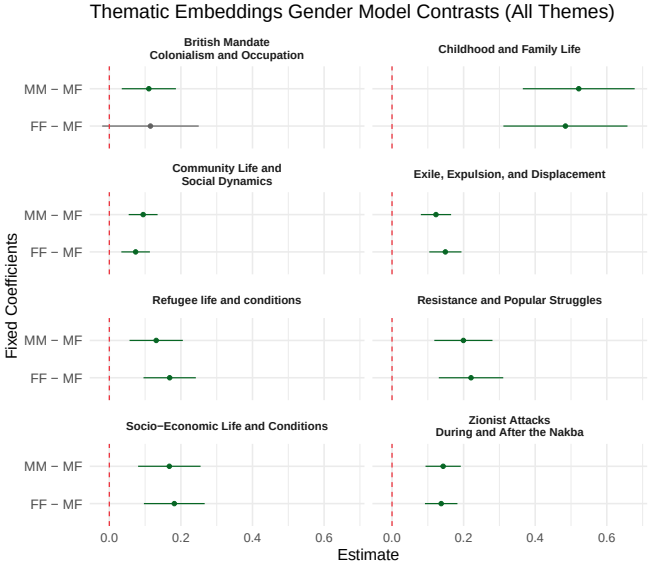


Figure 10. Mixed model for gender estimates for all themes, following the same description as in the main text.

Table 7. Comparison of model estimates for location contrasts across three modeling approaches.

Contrast	Mundlak (Bayes)	Fixed Effects	Mundlak (Freq.)
ar_families			
Both - None	2.136 [2.034, 2.239]*	2.240 [1.380, 3.099]*	2.241 [2.179, 2.304]*
OriginOnly - None	2.142 [2.097, 2.188]*	2.184 [1.644, 2.724]*	2.186 [2.133, 2.239]*
ResidenceOnly - None	0.015 [-0.005, 0.036]	0.007 [-0.187, 0.202]	0.007 [-0.043, 0.058]
Synergy	-0.021 [-0.134, 0.094]	0.048 [-0.620, 0.716]	0.048 [-0.015, 0.111]
ar_keywords			
Both - None	0.560 [0.478, 0.643]*	0.587 [0.325, 0.849]*	0.587 [0.565, 0.609]*
OriginOnly - None	0.205 [0.168, 0.242]*	0.243 [0.133, 0.354]*	0.244 [0.225, 0.263]*
ResidenceOnly - None	0.177 [0.161, 0.193]*	0.204 [0.114, 0.294]*	0.204 [0.186, 0.222]*
Synergy	0.178 [0.087, 0.270]*	0.139 [-0.081, 0.360]	0.139 [0.117, 0.161]*
ar_landmarks			
Both - None	0.806 [0.696, 0.916]*	0.946 [0.287, 1.606]*	0.948 [0.896, 0.999]*
OriginOnly - None	0.469 [0.420, 0.517]*	0.568 [0.224, 0.911]*	0.568 [0.524, 0.612]*
ResidenceOnly - None	0.145 [0.123, 0.166]*	0.136 [-0.039, 0.312]	0.136 [0.094, 0.178]*
Synergy	0.192 [0.073, 0.315]*	0.242 [-0.256, 0.741]	0.244 [0.192, 0.296]*
ar_significant_figures			
Both - None	2.556 [2.448, 2.663]*	2.807 [1.394, 4.221]*	2.809 [2.723, 2.895]*
OriginOnly - None	1.625 [1.577, 1.672]*	1.874 [1.294, 2.454]*	1.875 [1.802, 1.948]*
ResidenceOnly - None	0.030 [0.009, 0.051]*	0.038 [-0.193, 0.268]	0.037 [-0.033, 0.107]
Synergy	0.901 [0.784, 1.019]*	0.896 [-0.115, 1.907]	0.897 [0.810, 0.984]*
ner_texts			
Both - None	0.929 [0.847, 1.011]*	1.090 [0.763, 1.417]*	1.090 [1.066, 1.115]*
OriginOnly - None	0.636 [0.600, 0.672]*	0.746 [0.616, 0.875]*	0.746 [0.725, 0.767]*
ResidenceOnly - None	0.138 [0.122, 0.154]*	0.189 [0.101, 0.278]*	0.189 [0.170, 0.209]*
Synergy	0.154 [0.065, 0.244]*	0.155 [-0.058, 0.368]	0.155 [0.131, 0.180]*

Table 8. Comparison of model estimates for gender contrasts across three modeling approaches.

Contrast	Fixed Effects	Mundlak (Freq.)	Mundlak (Bayes)
ar_families			
FF - MF	0.036 [0.006, 0.067]*	0.037 [0.003, 0.071]*	NA
MM - MF	0.016 [-0.015, 0.048]	0.016 [-0.018, 0.050]	NA
ar_keywords			
FF - MF	0.088 [0.050, 0.126]*	0.088 [0.060, 0.116]*	0.109 [0.085, 0.133]*
FemaleFemale - MaleMale	NA	NA	0.028 [-0.015, 0.071]
MM - MF	0.103 [0.058, 0.147]*	0.103 [0.075, 0.131]*	0.081 [0.059, 0.103]*
ar_landmarks			
FF - MF	0.054 [0.020, 0.088]*	0.055 [0.017, 0.093]*	NA
MM - MF	0.045 [0.000, 0.090]*	0.045 [0.007, 0.083]*	NA
ar_significant_figures			
FF - MF	0.026 [0.004, 0.049]*	0.026 [-0.009, 0.062]	NA
MM - MF	0.029 [0.009, 0.049]*	0.029 [-0.006, 0.065]	NA
ner_texts			
FF - MF	0.130 [0.088, 0.173]*	0.130 [0.102, 0.158]*	NA
MM - MF	0.120 [0.076, 0.163]*	0.120 [0.092, 0.148]*	NA

Table 9. Comparison of model estimates for location contrasts across three modeling approaches. Estimates shown as: Estimate [95% CI]. * indicates non-zero effect.

Contrast	Mundlak (Bayes)	Fixed Effects	Mundlak (Freq.)
British Mandate Colonialism and Occupation			
Both - None	0.467 [0.284, 0.651]*	0.448 [0.099, 0.797]*	0.448 [0.400, 0.497]*
OriginOnly - None	0.011 [-0.082, 0.104]	0.051 [-0.128, 0.230]	0.051 [0.011, 0.091]*
ResidenceOnly - None	0.052 [0.012, 0.091]*	0.037 [-0.043, 0.117]	0.037 [0.002, 0.072]*
Synergy	0.405 [0.199, 0.612]*	0.360 [0.011, 0.709]*	0.360 [0.311, 0.409]*
Childhood and Family Life			
Both - None	0.417 [0.127, 0.711]*	0.319 [-0.023, 0.660]	0.320 [0.258, 0.382]*
OriginOnly - None	0.086 [-0.019, 0.191]	0.082 [-0.080, 0.243]	0.082 [0.046, 0.118]*
ResidenceOnly - None	0.074 [0.031, 0.116]*	0.082 [-0.006, 0.169]	0.082 [0.050, 0.113]*
Synergy	0.258 [-0.054, 0.576]	0.155 [-0.122, 0.433]	0.157 [0.097, 0.216]*
Community Life and Social Dynamics			
Both - None	0.122 [0.007, 0.236]*	0.073 [-0.117, 0.262]	0.073 [0.049, 0.097]*
OriginOnly - None	0.092 [0.045, 0.139]*	0.083 [0.007, 0.158]*	0.083 [0.064, 0.101]*
ResidenceOnly - None	0.048 [0.027, 0.068]*	0.060 [0.016, 0.104]*	0.060 [0.043, 0.077]*
Synergy	-0.018 [-0.143, 0.107]	-0.070 [-0.219, 0.079]	-0.070 [-0.093, -0.046]*
Exile, Expulsion, and Displacement			
Both - None	0.150 [0.044, 0.254]*	0.135 [0.022, 0.248]*	0.135 [0.115, 0.154]*
OriginOnly - None	0.021 [-0.018, 0.062]	0.015 [-0.045, 0.074]	0.015 [-0.000, 0.029]*
ResidenceOnly - None	0.011 [-0.007, 0.028]	-0.001 [-0.033, 0.030]	-0.001 [-0.015, 0.012]
Synergy	0.118 [0.003, 0.230]*	0.121 [0.023, 0.220]*	0.121 [0.102, 0.141]*
Refugee life and conditions			
Both - None	0.407 [0.228, 0.589]*	0.301 [0.032, 0.571]*	0.303 [0.255, 0.352]*
OriginOnly - None	0.111 [0.032, 0.191]*	0.065 [-0.054, 0.185]	0.066 [0.027, 0.104]*
ResidenceOnly - None	0.262 [0.226, 0.297]*	0.222 [0.117, 0.328]*	0.223 [0.188, 0.258]*
Synergy	0.035 [-0.163, 0.236]	0.014 [-0.193, 0.220]	0.014 [-0.031, 0.060]
Resistance and Popular Struggles			
Both - None	0.217 [0.088, 0.344]*	0.234 [-0.021, 0.489]	0.233 [0.203, 0.264]*
OriginOnly - None	0.007 [-0.050, 0.065]	-0.020 [-0.104, 0.065]	-0.019 [-0.045, 0.006]
ResidenceOnly - None	0.036 [0.013, 0.060]*	0.023 [-0.027, 0.073]	0.023 [0.001, 0.045]*
Synergy	0.173 [0.032, 0.313]*	0.231 [-0.004, 0.465]	0.230 [0.201, 0.259]*
Socio-Economic Life and Conditions			
Both - None	0.144 [0.029, 0.258]*	0.188 [-0.102, 0.478]	0.188 [0.152, 0.224]*
OriginOnly - None	0.044 [-0.017, 0.104]	0.074 [-0.051, 0.199]	0.074 [0.045, 0.104]*
ResidenceOnly - None	0.047 [0.019, 0.076]*	0.047 [-0.008, 0.103]	0.047 [0.019, 0.075]*
Synergy	0.053 [-0.079, 0.184]	0.067 [-0.238, 0.371]	0.067 [0.031, 0.102]*
Zionist Attacks During and After the Nakba			
Both - None	0.240 [0.160, 0.321]*	0.224 [0.045, 0.403]*	0.224 [0.206, 0.241]*
OriginOnly - None	0.106 [0.073, 0.139]*	0.087 [0.025, 0.150]*	0.087 [0.073, 0.102]*
ResidenceOnly - None	0.038 [0.023, 0.053]*	0.033 [-0.008, 0.075]	0.033 [0.020, 0.047]*
Synergy	0.096 [0.010, 0.184]*	0.103 [-0.063, 0.269]	0.103 [0.085, 0.120]*

Table 10. Comparison of model estimates for gender contrasts across three modeling approaches. Estimates shown as: Estimate [95% CI]. * indicates non-zero effect.

Contrast	Mundlak (Bayes)	Fixed Effects	Mundlak (Freq.)
British Mandate Colonialism and Occupation			
FF - MF	0.153 [0.083, 0.225]*	0.115 [-0.020, 0.250]	0.113 [0.037, 0.189]*
MM - MF	0.073 [0.014, 0.131]*	0.110 [0.035, 0.186]*	0.112 [0.044, 0.180]*
Childhood and Family Life			
FF - MF	0.600 [0.552, 0.647]*	0.484 [0.311, 0.658]*	0.484 [0.431, 0.538]*
MM - MF	0.408 [0.363, 0.452]*	0.522 [0.365, 0.678]*	0.522 [0.469, 0.575]*
Community Life and Social Dynamics			
FF - MF	0.096 [0.069, 0.123]*	0.073 [0.034, 0.113]*	0.073 [0.042, 0.104]*
MM - MF	0.073 [0.046, 0.099]*	0.094 [0.054, 0.135]*	0.095 [0.063, 0.126]*
Exile, Expulsion, and Displacement			
FF - MF	0.145 [0.121, 0.169]*	0.149 [0.104, 0.194]*	0.149 [0.123, 0.175]*
MM - MF	0.128 [0.105, 0.150]*	0.123 [0.081, 0.165]*	0.123 [0.097, 0.149]*
Refugee life and conditions			
FF - MF	0.180 [0.129, 0.232]*	0.169 [0.095, 0.242]*	0.167 [0.107, 0.228]*
MM - MF	0.121 [0.071, 0.170]*	0.131 [0.057, 0.205]*	0.132 [0.072, 0.193]*
Resistance and Popular Struggles			
FF - MF	0.265 [0.229, 0.300]*	0.221 [0.131, 0.311]*	0.221 [0.182, 0.259]*
MM - MF	0.156 [0.123, 0.188]*	0.199 [0.118, 0.281]*	0.199 [0.162, 0.237]*
Socio-Economic Life and Conditions			
FF - MF	0.216 [0.174, 0.258]*	0.182 [0.097, 0.266]*	0.181 [0.134, 0.228]*
MM - MF	0.134 [0.095, 0.173]*	0.168 [0.080, 0.255]*	0.168 [0.122, 0.215]*
Zionist Attacks During and After the Nakba			
FF - MF	0.160 [0.140, 0.180]*	0.137 [0.092, 0.183]*	0.137 [0.115, 0.160]*
MM - MF	0.121 [0.102, 0.140]*	0.143 [0.094, 0.192]*	0.143 [0.121, 0.165]*

Embedding similarity insights

POHA contains rich information about experiences in different parts of pre-1948 Palestine. Here, we present an additional table of contents-based plots. Figure 11 reports the average thematic similarity between interview segments for every place of origin represented in POHA.

We also explore the co-occurrence of different themes, reported in fig. 12. The theme scheme we use engenders significant overlap among several categories namely, Zionist attacks, expulsion, and resistance.

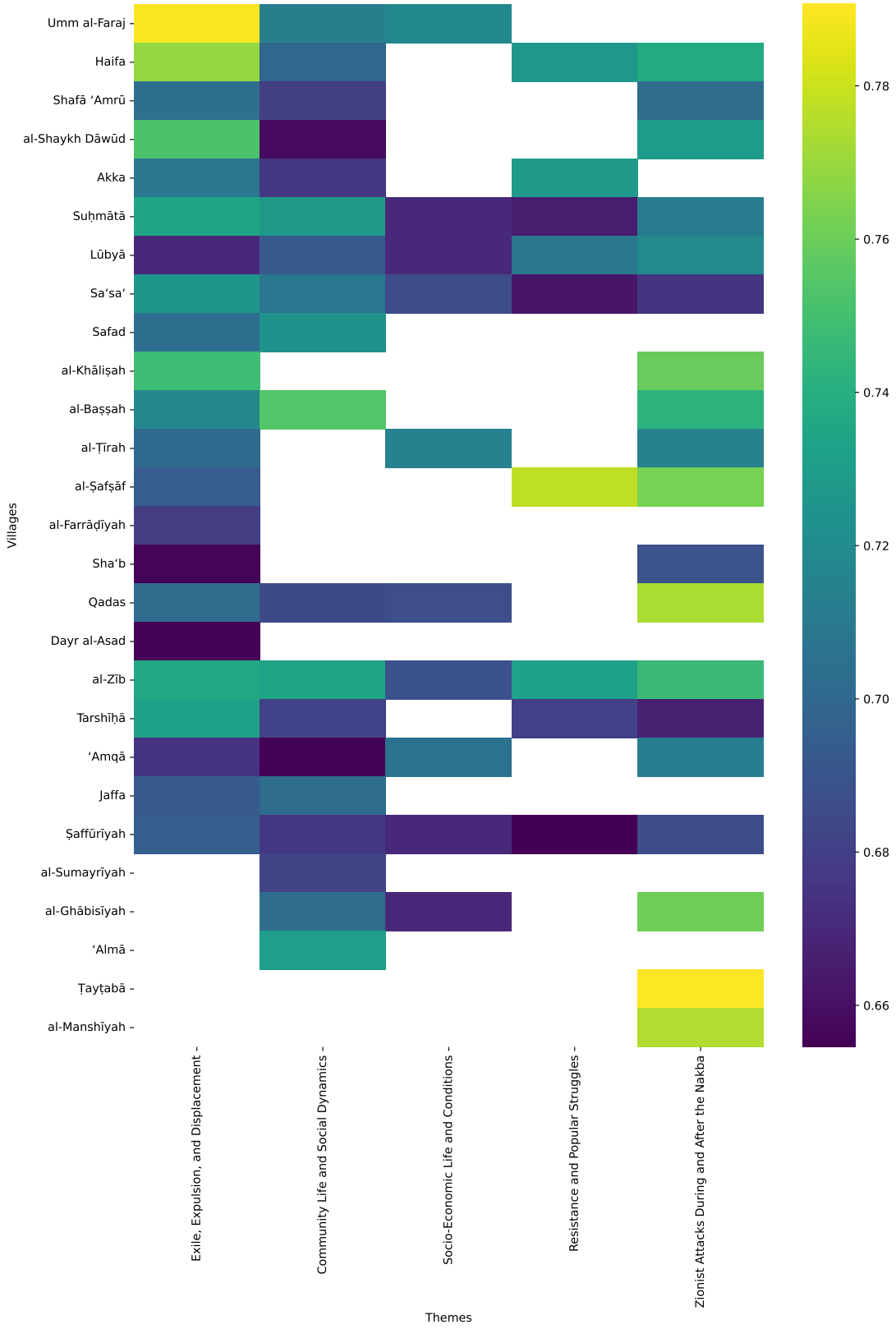


Figure 11. Average cosine similarity between the embeddings of two interview segments on the labeled theme for every place of origin in POHA. A blank rectangle represents a lack of data.

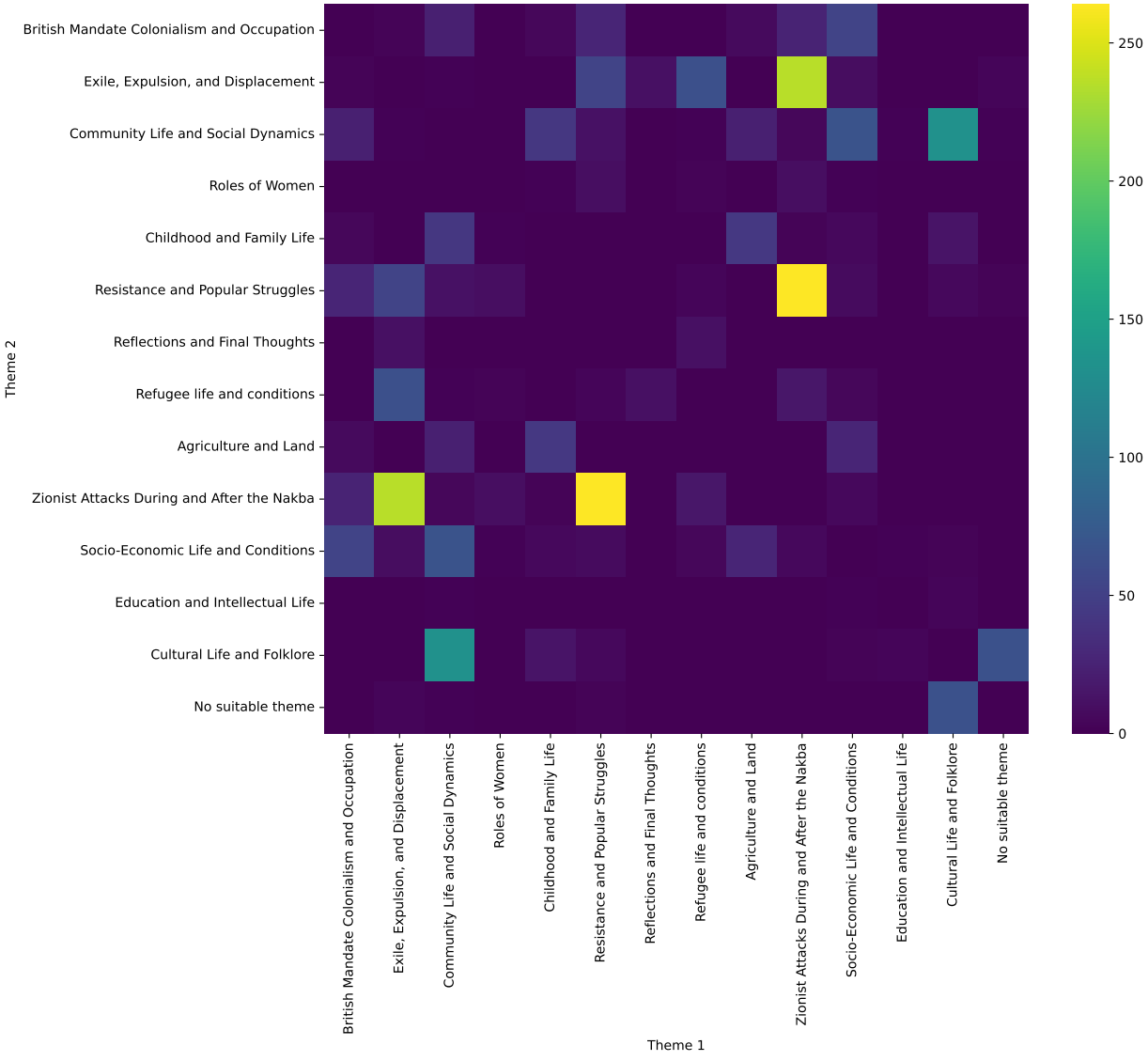


Figure 12. Cooccurrence between themes in POHA interview segments (i.e., the number of headers for which one theme and another both apply).