

## 재난 피해자의 지속적 어려움과 심리적 지원 필요성: 면담 텍스트 기반 발화 분석

김 지 호<sup>1,2</sup> · 이 소 희<sup>1</sup> · 임 혜 선<sup>1</sup> · 남 동 엽<sup>1</sup> · 황 우 석<sup>1</sup> · 양 수 연<sup>1</sup> · 박 상 현<sup>1\*</sup>

<sup>1</sup>국립재난안전연구원 재난원인조사실 재난회복연구센터

<sup>2</sup>부산대학교 정보융합공학과 박사과정

## Persistent Difficulties and the Need for Psychological Support Among Disaster Victims: A Speech-Based Analysis of Interview Texts

Ji-Ho Kim<sup>1,2</sup> · So-Hee Lee<sup>1</sup> · Hye-Sun Lim<sup>1</sup> · Dong-Yup Nam<sup>1</sup> · Woo-Suk Hwang<sup>1</sup> · Su-Yeon Yang<sup>1</sup> · Sang-Hyun Park<sup>1\*</sup>

<sup>1</sup>Disaster Prevention Research Division, National Disaster Management Research Institute, Ulsan 44538, Korea

<sup>2</sup>Department of Information Convergence Engineering, School of Computer Science and Engineering, Pusan National University, Busan 46241, Korea

### [요 약]

재난은 발생 이후에도 장기간에 걸쳐 개인의 삶에 영향을 미치며, 이러한 영향은 피해자가 자신의 상태와 경험을 서술하는 방식에 반영된다. 본 연구는 면담 텍스트를 기반으로 표현의 분포와 결합 양상을 분석하고, 이를 통해 재난 이후 경험이 어떠한 형태로 나타나는지를 확인하고자 하였다. 특히 사전 정의된 감정어에 의존하지 않고 면담 과정에서 생성된 표현을 분석 대상으로 설정하여, 언어적 서술 양상을 탐색하였다. 분석 결과 감정표현은 개별 감정어로 분리되어 나타나기보다, “그 이후에”, “잠도 못 자고”와 같이 현재 상태와 사건 이후의 경험이 결합된 형태로 반복적으로 나타나는 경향이 확인되었다. 또한 반복 표현과 지원 체계 관련 표현이 함께 나타난다는 점에서 재난 이후 경험이 현재 상태와 연결된 형태로 서술되는 양상이 관찰되었으며, 유가족 집단은 평균적으로 보다 다양한 표현을 사용하였고, 생존자 집단은 개인 간 편차가 크게 나타났다. 이러한 결과는 재난 이후 경험이 집단에 따라 서로 다른 방식으로 구성되고 표현될 수 있음을 보여준다. 본 연구는 재난피해자의 언어적 특성을 이해하기 위한 기초적 근거를 제공하고 장기적인 관점에서의 심리적 지원 제도 마련이 필요함을 시사한다.

### [Abstract]

Disasters exert long-term impacts on individuals even after the initial event, and these impacts are reflected in how victims describe their conditions and experiences. This study analyzes the distribution and co-occurrence patterns of expressions in interview texts to examine how post-disaster experiences manifest. Rather than relying on predefined emotion lexicons, the study utilizes expressions generated during the interviews, allowing for an empirical exploration of linguistic characteristics. The results show emotional expressions do not appear as isolated words, instead they manifest as recurring combinations linking current conditions and post-event experiences, such as “since then” and “unable to sleep.” The expression pattern or recurring expression combinations and references to support systems further indicates that post-disaster experiences are described in connection with present conditions. The bereaved group showed a greater diversity of expressions on average, while the survivor group showed greater individual variation in expression use. These findings suggest that post-disaster experiences are constructed and articulated differently across groups. This study provides empirical evidence for understanding the linguistic characteristics of post-disaster experiences and suggests the need for the need for long-term psychological support.

**색인어** : 재난 회복, 재난 이후 경험, 면담 텍스트 분석, 심리적 지원, 텍스트 마이닝

**Keyword** : Disaster Resilience, Post-Disaster Experience, Interview Text Analysis, Psychological Support, Text Mining

<http://dx.doi.org/10.9728/dcs.2026.27.5.1323>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Received** 19 April 2026; **Revised** 06 May 2026

**Accepted** 14 May 2026

**\*Corresponding Author, Sang-Hyun Park**

**E-mail:** mavin76@korea.kr

## 1. 서론

재난은 발생 이후 장기간에 걸쳐 개인의 일상과 사회적 관계에 영향을 미치는 사건이다. 이 과정에서 생존자와 유가족은 재난 경험을 반복적으로 회상하고 설명하며 자신의 상태와 변화를 언어로 표현한다. 이러한 언어적 표현은 재난이 개인의 삶 속에서 어떻게 인식되고 재구성되는지를 보여주는 중요한 자료로 작용한다. 따라서 재난 이후 회복을 이해하기 위해서는 피해자의 경험이 축적된 언어적 기록에 주목할 필요가 있다.

재난 이후 피해자의 발화는 동일한 사건을 설명하더라도 표현 방식에서 차이를 보인다. 재난을 경험한 개인들은 특정 표현을 반복적으로 사용하여 재난피해 경험을 설명한다. 반면, 동일한 재난을 경험한 또 다른 피해자는 다양한 어휘를 사용하여 상황을 서술한다. 이러한 차이는 발화의 길이, 사용되는 표현의 반복성, 감정 관련 표현의 밀도 등 언어적 양상으로 관찰될 수 있다. 이는 재난 이후 개인이 느낀 감정과 경험이 회복 과정 중 다른 형태로 진행되기 때문이다[1].

기존 연구에서는 재난 이후 피해자의 경험을 면담이나 사례 분석을 통해 해석하는 방식이 주로 사용되어 왔다[2]. 이러한 접근은 개별 사례의 맥락을 이해하는데 유용하지만 다수의 피해자를 대상으로 나타나는 표현 양상의 차이를 체계적으로 비교하는 데에는 한계가 있다. 특히 피해자의 발화에 포함된 반복 표현이나 감정 관련 어휘의 사용양상은 정성적 분석만으로는 일관되게 파악하기에 한계가 있다.

최근에는 뉴스 기사나 온라인 텍스트를 대상으로 텍스트 마이닝 기법을 적용하여 재난 관련 이슈와 사회적 반응을 분석하는 연구가 증가하고 있다. 이러한 연구들은 특정 키워드의 빈도나 주제 분포를 통해 재난 담론의 구조를 파악하고자 하였다[3],[4]. 그러나 실제 피해자의 면담 텍스트를 대상으로 언어적 표현의 구조 자체를 분석한 연구는 상대적으로 제한적이다.

재난 피해자의 면담 텍스트에는 불안, 두려움, 답답함, 분노와 같은 다양한 감정 표현이 혼재되어 나타날 수 있다. 이러한 표현은 특정한 감정 단어로만 구성되지 않으며, “잠을 못 잔다.,” “생각이 난다.” 등과 같이 일상적 표현의 형태로 드러나는 경우가 존재한다. 따라서 감정 표현을 사전 기반으로 단순 분류하는 방식은 피해자의 언어를 충분히 반영하기 어렵다.

본 연구에서는 재난 피해자 면담 텍스트에 나타난 감정 관련 표현을 사전에 의존하지 않고 n-gram 기반으로 추출한다. 이를 통해 피해자가 실제로 사용한 표현단위를 중심으로 감정 표현 후보를 수집하고, 개인별로 어떠한 표현이 반복적으로 사용되는지를 분석한다. 이러한 접근은 감정 표현을 미리 정의된 범주에 맞추기보다 피해자의 언어에서 자연스럽게 드러나는 표현들을 구조적으로 탐색한다.

피해자별 감정 표현의 사용 양상은 표현의 개수, 종류, 반복성 측면에서 차이를 보인다. 본 연구에서는 개인 별로 사용

된 감정 표현의 다양성을 엔트로피 지표로 산출하고 표현 분포의 유사성을 코사인 유사도로 비교한다. 이를 통해 특정 피해자가 제한된 표현을 반복적으로 사용하는지 혹은 다양한 표현을 사용하는지를 정량적으로 파악할 수 있다. 이러한 분석은 회복 과정에서 나타나는 언어적 차이를 구조적으로 비교하는 기반을 제공한다.

본 연구는 재난 피해자의 심리 상태를 진단하거나 분류하는 것을 목적으로 하지 않는다. 피해자들은 재난 이후 경험을 어떻게 언어로 표현하는지에 대해 집중하고 이러한 표현의 구조적 특성을 분석하고자 한다. 언어는 개인의 상태를 직접 표현하는 지표를 넘어 회복 과정에서 나타나는 변화가 축적된 결과로 이해한다.

본 연구는 재난 이후 생존자와 유가족을 대상으로 수집된 면담 텍스트를 분석 대상으로 설정하고, 재난 이후 경험이 실제 발화에서 어떠한 방식으로 구성되고 서술되는지를 정량적으로 확인하고자 하였다. 이를 위해 면담자료로부터 사전 정의에 의존하지 않고 n-gram 기반 방식으로 감정표현 후보를 추출하고, 표현 간 공출현 관계를 분석하여 감정표현의 결합 양상을 확인한다. 또한 개인 및 집단 단위에서 나타나는 감정 표현의 사용량과 다양성 및 유사성을 비교함으로써 재난 이후 경험이 언어적으로 나타나는 차이를 확인한다. 이를 통해 재난 이후 경험이 개인과 집단에 따라 어떠한 방식으로 지속되고 표현되는지를 실증적으로 제시하고, 장기적인 회복 과정에서 나타나는 어려움과 지원 필요성에 대한 기초적 근거를 제시하고자 한다.

## II. 관련 연구

### 2-1 재난 회복 과정에 관한 연구

최근 재난 연구에서 회복은 물리적 피해의 복구를 넘어 재난 피해자가 일상생활로 복귀하고 사회적, 경제적 맥락 속에서 적응하고, 개선하고 변화하는 능력으로 정의된다[1]. 이러한 관점에서 재난에 대한 회복은 단기적인 복구 성과가 아니라 시간의 경과에 따라 축적되는 경험과 적응의 결과로 이해될 수 있다.

재난 회복을 가능하게 하는 핵심 개념으로 회복탄력성이 논의되어 왔다. 회복탄력성은 다양한 정의가 존재하지만 외부 교란에도 생태계가 원래 상태를 유지할 수 있는 능력으로 나타났다[5]. 해당 용어는 이후 개인, 인간 공동체 및 사회의 적응 능력을 설명하는데 사용되었다[6]. 이에 회복탄력성은 외부 교란이나 스트레스에도 안정적인 시스템을 유지하는 능력, 대규모 장애 발생 후 복구하는 능력 그리고 새로운 환경에 적응하는 능력이 핵심이다.

재난 회복에 영향을 미치는 요소는 크게 외적 요인과 내적 요인으로 구분할 수 있다[7],[8]. 외적 요인은 제도적 지원, 물리적 환경, 자원 접근성, 사회적 인프라와 같이 개인이 통제하기 어려운 조건을 포함한다. 반면 내적 요인은 개인이 경험

하는 신체적, 심리적 상태, 일상 수행 능력, 주거 및 생활 환경에 대한 인식과 같이 개인의 삶의 질과 직접적으로 연관된 요소로 정리된다.

특히 장기적인 회복 과정에서 내적 요인은 피해자의 삶의 질과 기능 회복 여부에 큰 영향을 미친다. 재난 이후 일정 시간이 경과하더라도 피해자는 불면, 통증, 불안, 무기력 등과 같은 신체적이고 심리적인 고통을 지속적으로 경험하는 경우가 많다[9]. 이러한 고통은 단순한 개인적 문제를 넘어 사회적 관계 유지, 일상생활, 생계유지 등을 제약하게 되며, 회복 과정 전반을 지연시키는 요인으로 작동한다.

그럼에도 불구하고 기존 재난 회복 연구에서는 회복상태를 정량적으로 평가하기 위해 주로 인프라 복구 수준, 제도적 대응, 주거 안정성과 같은 물리적이고 구조적인 지표를 우선적으로 활용하였다. 이러한 접근은 재난 대응의 가시적 성과를 비교하고 평가하는데 유용하다. 하지만 그 과정에서 피해자가 실제로 경험하는 신체적이고 심리적인 고통에 대한 미시적인 회복과정은 구체적인 분석 대상에서 제외되는 경향이 나타났다. 특히 재난 회복은 피해자의 내적 경험을 충분히 반영하여 장기적인 관점에서 회복상태를 고려해야 하나 상대적으로 충분히 다루어지지 못한 한계가 있다.

이러한 맥락에서 재난 회복을 정밀하게 이해하기 위해서 피해자의 주관적 경험과 일상적 고통이 어떻게 축적되고 표현되는지를 분석할 필요가 있다. 재난 피해자가 사용하는 언어는 회복 과정에서의 상태를 반영하는 중요한 단서가 될 수 있으며, 이것은 기존의 정량 지표로는 포착하기 어려운 회복의 질적 측면을 확인할 수 있다. 따라서 재난 회복 연구에서 텍스트 기반 언어분석은 회복력의 내적 차원을 규명하기 위한 중요한 방법론적 접근이다.

## 2-2 재난 연구에서의 텍스트 기반 언어 분석 연구

재난 회복 연구에서 개인의 내적 경험을 이해하기 위한 방법으로 텍스트 데이터의 활용이 점차 증가하고 있다. 재난 이후의 회복은 개인의 심리적 상태와 신체적 부담, 일상 수행의 변화가 복합적으로 얽힌 과정으로 나타난다. 이러한 경험은 표준화된 지표만으로 충분히 포착하기 어렵다. 이에 따라 재난을 경험한 사람이 자신의 경험을 서술하는 면담 기록과 자유 응답 텍스트 자료는 내적 요인을 탐색하는 중요한 자료이다.

기존의 재난 연구에서는 설문조사나 수치화된 지표를 통해 회복 상태를 비교하고 평가하는 접근이 주로 활용되어 왔다. 이러한 방법은 집단 간 비교와 정책적 판단에 유용한 정보를 제공하지만, 재난 피해자가 경험하는 감정의 변화나 일상 속 어려움과 같은 미시적 경험을 충분히 반영하는 데에는 한계가 있다. 동일한 피해나 경험을 한 사람들도 개인별 회복양상이 상이하며 이러한 차이는 개인의 언어 표현을 통해 직접적으로 드러난다.

언어는 재난 피해자가 자신의 경험을 해석하고 의미화하는 과정이 자연스럽게 반영된 매개체이다[10]. 면담과 서술형

응답에서 나타나는 표현은 특정 시점의 상태뿐 아니라, 재난 이후 시간의 흐름에 따라 누적되는 감정과 인식의 변화를 함께 포함한다. 이러한 특성으로 인해 텍스트 자료는 재난 회복을 결과가 아닌 과정으로 이해하려는 연구 흐름에 적합하다.

최근 연구에서는 키워드 빈도 분석이나 주제 분류와 같은 정량적 기법을 활용하여 대규모 텍스트 자료를 분석하려는 시도가 이루어지고 있다. 이러한 접근은 재난 이후 생산되는 방대한 텍스트를 요약하고 반복적으로 등장하는 이슈나 관심 주제를 파악하는 연구로 진행되었다. 특히 분석 과정의 일관성과 재현성을 확보하기 위해서 다양한 텍스트 분석 방법론이 재난 연구에서 점진적으로 적용되었다[11],[12].

텍스트 마이닝 연구에서는 어떠한 이슈와 개념이 반복적으로 언급되는지를 파악하기 위해 단어 빈도 기반 분석과 공출현(Co-occurrence word) 기반 분석이 활용되어 왔다. 공출현 분석은 동일한 문맥 내에서 함께 등장하는 단어들의 관계를 분석하는 방법이다. 개별 단어의 빈도만으로는 파악하기 어려운 의미적 맥락을 탐색하기 위해 사용된다. 공출현은 특정 경험이나 감정이 어떠한 상황이나 사건과 함께 서술되는지를 파악하는 경우에 사용한다[13]. 빈도 기반 분석은 단어 또는 표현의 출현 빈도를 집계함으로써, 사건 이후 텍스트에서 상대적으로 강조되는 핵심 개념이나 관심사를 식별한다. 이러한 접근은 재난 피해 및 복구 관련 대규모 텍스트 자료에서 주요 이슈를 개괄적으로 파악하는 것이 가능하다[14].

또한, 텍스트에 내재된 주요 논의 흐름을 체계적으로 파악하기 위해 잠재 디리클레 할당(Latent Dirichlet Allocation, LDA)을 활용한 주제 모델링 연구가 수행되어 왔다. LDA는 문서를 여러 주제의 확률 분포로 표현하는 분석 기법으로 대규모 텍스트 자료에 포함된 잠재적 주제 구조를 탐색하는데 사용된다. 기존 연구는 LDA를 활용하여 트위터에서 홍수 발생 기간 동안 논의된 주제를 식별하였다[15].

다른 방법으로 텍스트 간 의미적 유사성을 분석하기 위해 임베딩 기반 분석이 사용된다. 단어 임베딩이나 문서 임베딩은 텍스트를 벡터 공간에 표현하여 단어 또는 문서 간 의미적 거리를 수치적으로 계산할 수 있도록 한다. 이러한 임베딩 기반 접근은 면담자료에서 언어적으로 어떻게, 얼마나 유사한지 분석하는 데 사용된다. 이에 재난 관련 문맥을 이해하기 위해서 BERT 임베딩을 활용하여 재난 소셜 미디어에서의 재난을 예측한 연구가 수행된 바가 있다[16].

군집분석은 임베딩 기법을 기반으로 결합되어 텍스트 문서의 유형을 분류하는 연구로 진행된다. 군집분석은 텍스트를 임베딩 기법으로 벡터 공간에 매핑하고 벡터들 간 유사성을 기반으로 문서나 발화를 여러 집단으로 구분하는 클러스터링 등 비지도 학습 기법으로 사전에 범주를 정의하지 않고 반복적으로 나타나는 서술 유형을 탐색할 수 있다는 특징을 지닌다[17].

한편, 일부 연구에서는 재난 이후 나타나는 정서적 반응과 심리적 상태를 이해하기 위해 텍스트에 포함된 정서 표현에 주목하였다. 이러한 연구는 재난 피해자와 관련 경험자와의

텍스트를 면밀하게 읽고 그 안에 나타난 감정이나 정서적 단서를 질적으로 분석함으로써 재난 경험이 개인의 상태에 미치는 영향을 탐색하였다. 이에 쓰나미, 원자력 발전소 사고 등이 일어난 일본 도호쿠 지역에 방문한 해외 관광객을 대상으로 반구조화된 이메일 인터뷰를 통해 정서적 특성과 감정을 분석한 연구가 존재한다[18]. 이러한 해석적 접근을 통해 불안, 공포, 스트레스 등과 같은 정서적 반응이 어떻게 표현되는지를 확인하였다. 다만, 이러한 논의는 주로 정서의 존재와 의미를 해석하는데 초점을 두었으며 언어 표현 자체의 차이나 구조를 체계적이고 정량적으로 비교하는 단계로까지는 확장되지 못했다는 한계가 있다.

이에 따라, 기존의 텍스트 기반 접근에는 두 가지 한계가 존재한다. 첫째는 다수의 정량적 텍스트 데이터 마이닝 연구에서 언론 보도나 온라인 게시글과 같은 간접적으로 생성된 텍스트를 주요 분석 대상으로 사용하였다. 이러한 자료는 재난에 대한 사회적 반응을 파악하는 것에 유용하나 재난 피해자가 자신의 경험을 직접 구성하고 서술하는 언어의 특성을 충분히 반영하기는 어렵다. 이로 인해 회복 과정의 개인별 차이를 언어 수준에서 분석하는 것에는 제약이 있다. 두 번째는 정서 표현을 중심으로 한 질적 해석 연구는 언어적 표현의 비교나 집단 간 양상을 구조적으로 설명하지 못한 한계가 있다. 질적 연구는 재난 이후의 심리적 상태를 이해하는데 중요한 논의를 제공하나 이러한 접근은 개별 사례의 해석에 머무르는 경우가 많고 언어적 이질성과 패턴을 종합적으로 설명하기 어렵다.

결과적으로 본 연구는 이러한 한계를 보완하기 위해 재난 피해자의 직접적인 서술이 담긴 면담 텍스트를 분석 대상으로 설정하였다. 기존 텍스트 기반 재난 연구가 감정의 분류나 주제 파악에 집중하여 표현이 갖는 의미를 해석하였으나, 본 연구는 감정 표현의 분포와 결합 양상을 나타낸다. 또한 생존자와 유가족을 구분함으로써 집단 간 차이를 구체적으로 확인하고자 하였다. 또한, 집단 간 감정표현 사용 양상의 차이를 통해 동일한 재난 경험이 상이한 방식으로 언어화되고 있음을 확인하였다. 이에 재난 피해자의 발화를 기반으로 감정표현의 분포와 사용 양상을 함께 분석하고 표현이 나타나는 방식과 의미를 정량적으로 제시하였다는 점에서 의의가 있다.

### III. 연구 방법

#### 3-1 연구 자료

본 연구의 분석 대상은 2023년 충청북도 청주시 오송읍의 궁평2지하차도에서 발생한 재난 사건 이후 재난 피해를 경험한 생존자 및 유가족을 대상으로 수행된 21명의 면담 텍스트 데이터이다. 본 연구는 피해자의 익명성과 인권 보호를 고려하여 분석 과정에서 연구 참여자의 식별 가능성을 최소화하고 연구 결과가 개별 사건이나 인물로 환원되는 것을 방지한다.

면담은 음성으로 기록된 후 전문 속기사를 통해 전사되어 텍스트 자료로 변환되었다. 전사 과정에서는 발화의 순서와 내용을 원자료와 동일하게 유지하였다. 모든 면담 참여자는 연구 참여와 자료 활용에 대해 사전 동의를 제공하였다. 전사된 텍스트는 분석 단계에서 개인을 식별할 수 있는 정보가 제거된 상태로 활용되었다. 이와 같은 절차를 통해 연구 윤리를 충족하는 분석 자료를 구축하였다.

분석에 활용된 자료는 연구 목적에 따라 일부 사례가 아닌 동일한 기준으로 수집된 면담 자료 전체를 대상으로 한다. 본 연구는 면담에 참여한 모든 대상자의 발화를 포함하였으며 이를 통해 재난 피해 경험이 지니는 언어적 표현의 다양성과 분포를 집단 수준에서 파악한다.

면담은 사전에 구성된 질문 흐름을 기반으로 진행되었다. 면담 질문은 피해 경험의 시간적 전개와 삶의 영역별 변화를 중심으로 설계되었다. 면담 참여자는 자신의 경험을 서사적으로 설명할 수 있도록 유도되었다. 질문은 면담 참여의 동기 및 초기 인식에서 시작하여 재난 이후 초기 수습 과정과 이후 일상 생활에서의 변화 그리고 현재 인식하는 회복 상태와 어려움, 마지막으로 향후 필요한 지원과 제도적 개선에 대한 인식으로 질의하였다.

#### 3-2 데이터 전처리

본 연구의 데이터 전처리는 재난 면담 텍스트에서 나타나는 표현 양상을 체계적으로 분석하기 위해서 면담 발화를 분석 가능한 단위로 변환하는 전처리 과정을 수행하였다. 전처리 과정은 면담 발화의 내재된 언어적 특성을 표현 기반 분석에 활용할 수 있도록 데이터를 설정하였다. 데이터의 보존성과 분석 효율성을 확보하기 위해 문장 데이터는 원문 문장과 분석용 문장으로 분리하여 관리하였다. 원문 문장은 면담 참여자의 실제 발화를 그대로 보존하여 인용과 해석 검토에 활용할 수 있도록 하였다. 분석용 문장은 표현 매칭과 빈도 집계를 안정적으로 수행하기 위해 토큰화와 기호 제거를 중심으로 정규화한 것이다.

구체적으로 면담 발화의 주체를 구분하기 위해 화자 표식을 정규식 기반으로 처리하였다. 화자 표식은 발화 주체를 식별하기 위한 정보로 활용하였으며 해당 정보는 화자 역할 변수로 분리하여 저장하였다. 발화 내용에는 화자 표식 문자열을 포함하지 않고 실제 발화 텍스트만 유지하였다. 이러한 처리는 조사자의 질문이나 안내 발화와 면담 참여자의 경험 서술을 구분하여 분석 대상 발화를 명확히 설정하기 위한 것이다. 이후 모든 분석 단계에서는 면담 참여자의 발화만을 사용하였다.

재난 면담에서 감정과 불안은 단일 단어보다 여러 단어가 함께 사용되는 표현 형태로 나타나는 경우가 많다. 이러한 표현을 분석하기 위해서 발화를 일정한 단위로 구분하고 각 단위 안에서 함께 사용된 단어 묶음을 확인하고자 하였다. 이에 본 연구에서는 재난 면담 발화를 문장 기준으로 분리하였다.

문장은 표현이 함께 사용되는 문맥 단위로 설정되었으며, 이후 n-gram 기반 표현 후보 수집과 표현 간 관계 분석은 모두 이러한 단위를 기준으로 수행하였다.

각 문장에는 그룹, 면담자 식별자, 발화 및 문장 식별자를 부여하였다. 그룹은 면담 참여자의 유형을 기준으로 구분되는 범주로 생존자 그룹과 유기족 그룹으로 구성된다. 면담자 식별자는 각 인터뷰 대상자에 대한 식별자이며 발화 식별자는 한 사람이 연속해서 응답한 발화 단락을 의미한다. 문장 식별자는 여러 문장으로 이루어진 발화를 문장 단위로 구분한 것이다. 이러한 식별자는 분석 결과를 문서와 발화 수준으로 추적할 수 있도록 구성하였다.

이와 같은 전처리 과정을 통해 구축된 문장 단위 데이터셋은 표현 후보 수집, 표현 사전 구축, 표현 공출현 분석 등 분석의 입력 자료로 활용되었다. 본 연구의 데이터 전처리기는 면담에서 사용된 언어 표현을 문장 단위로 정렬하고 경험 유형에 따라 나타나는 표현 양상을 분석할 수 있도록 구성하였다.

### 3-3 분석 방법

본 연구는 재난 면담 텍스트에서 감정과 불안 등의 표현이 어떤 형태로 반복되고 결합되는지를 파악하는 것을 목표로 분석하였다. 이를 위해 면담 참여자의 발화를 대상으로 문맥 단위 안에서 사용된 어절 묶음을 추출하고 해당 표현들이 어떤 방식으로 함께 사용되는지를 단계적으로 분석하였다.

분석은 면담 참여자의 발화 텍스트를 사용하여 전처리 단계에서 화자 구분이 완료된 문장 데이터를 입력으로 사용하였다. 문장은 표현이 함께 사용되는 문맥 범위를 설정하는 기준이다. 실제 분석 대상은 각 문장 내부에서 생성된 어절 단위 표현으로, 문장 간 결합을 통한 표현 생성은 수행하지 않았다. 이를 통해 표현이 어느 문맥 안에서 사용되는지를 문장 수준에서 파악하고자 하였다.

연구는 면담 텍스트에 포함된 감정과 불안 관련 표현을 어절 단위의 조합과 사용 맥락에서 파악하고자 분석을 설계하였다. 이를 위해 분석 대상 표현을 단일 어절이 아니라, 문장 안에서 함께 사용되는 어절들의 결합 형태로 분석을 적용하고자 하였다. 이러한 분석 목적에 따라 문장 내부에서 연속적으로 사용된 어절 조합을 체계적으로 수집할 수 있는 n-gram 기반 분석을 적용하였다.

n-gram 분석은 텍스트 내에서 연속적으로 사용된 어절의 결합을 추출하는 방법이다. 이러한 분석은 말의 사용 양상을 단어 조합 수준에서 문맥의 형태로 파악한다. 재난 면담 텍스트에 나타난 표현은 문장 안에서 함께 사용되는 어절의 결합 형태를 포착한다. 이를 통해 면담 참여자가 자신의 경험과 상태를 서술하는 과정에서 반복적으로 사용되는 말의 행태를 수집한다.

n-gram의 생성은 면담 발화에서 반복적으로 사용된 어절 결합 형태를 수집하기 위한 절차로 수행된다. 전처리 단계에서 각 문장을 대상으로 어절 단위 토큰을 생성하고, 인접한

어절의 연속 조합을 산출하였다. 분석에서는 2~5개의 토큰 조합으로 분석을 수행하였다. 이와 같이 수집된 표현 후보 목록은 검토하여 분석 목적에 부합하는 표현을 선별하여 표현 사전으로 구성한다.

표현 사전은 n-gram 분석을 통해 수집된 표현 후보 중에서 분석 목적에 부합하는 표현을 기준 목록으로 정리한 것이다. 본 논문에서는 자동으로 수집된 표현 후보 목록을 검토하여 면담 참여자의 상태나 경험을 드러내는 어절 결합을 중심으로 표현을 선별하였다. 이 과정을 통해 분석 대상 표현을 명확히 설정하고 이후 단계에서 일관된 기준으로 표현을 식별할 수 있도록 하였다.

구축된 표현 사전은 각 문장에서 사용된 표현을 확인하는 매칭 절차를 수행하였다. 매칭은 문장 단위를 기준으로 이루어졌으며 하나의 문장에 여러 표현이 동시에 기록될 수 있도록 구성하였다. 이를 통해 면담 발화에서 표현이 단독으로 사용되는 경우와 여러 표현이 함께 사용되는 양상을 함께 기록하도록 구성하였다. 표현 매칭 결과는 각 문장 식별자를 기준으로 정리되어 이후 분석의 입력자료로 활용한다.

표현 간 사용 양상의 비교를 위해 문장과 표현의 대응 관계를 행렬로 구성하였다. 행은 문장 번호 식별자를 사용하고 열은 표현 사전의 어절을 사용하여 희소 행렬을 구성한다. 이를 통해 각 문장에서의 표현이 어떤 문맥에서 사용되었는지를 정량적으로 정리한다.

공출현 분석은 동일한 문장 안에서 함께 등장한 표현 쌍을 기준으로 수행한다. 이를 통해 특정 표현들이 어떤 조합으로 함께 사용되는지를 확인한다. 반복적으로 함께 나타나는 표현 조합은 표현 간 사용 맥락의 유사성을 파악하며 이후 표현 간 관계를 비교한다.

공출현 결과를 바탕으로 표현 간 사용 맥락의 유사도를 산출하였다. 유사도 지표로는 자카드 유사도와 PMI(Pointwise Mutual Information)를 사용하였다. 자카드 유사도는 두 표현이 함께 등장한 문장 수를, 두 표현 중 하나라도 등장한 전체 문장 수로 나눈 값으로 정의된다. 해당 지표를 통해 표현들이 동일한 발화 맥락에서 함께 사용되는 정도를 비교할 수 있으며 반복적으로 함께 사용되는 표현 조합을 식별한다. PMI는 두 표현이 함께 등장한 확률이 각 표현이 독립적으로 등장할 확률의 곱으로 나눈 로그값이다. 이를 통해 각 표현의 등장 빈도를 고려하여 해당 표현이 결합된 표현인지를 확인한다.

본 논문에서는 Python 3.10 환경에서 pandas 및 numpy를 활용한 행렬 연산 및 집계를 수행하였으며, 네트워크 기반 군집화를 위해 Networkx 3.4.2 버전을 사용하였다.

## IV. 분석 결과

### 4-1 공출현 기반 감정표현 결합 분석의 양상

전체 면담자료는 총 18,245개의 문장으로 구성되었으며,

문장당 평균 9.67개의 어절과 27.15자의 발화 길이를 나타내었다. 본 연구는 참여자의 발화만을 대상으로 감정표현 분석을 수행하였다. 감정표현 후보는 각 문장에서 인접한 2~5개의 연속 어절 결합(n-gram)을 기반으로 감정표현 후보를 추출하였다. 이 과정에서 접속어, 대명사, 주입사, 일반 명사 등을 불용어로 처리하였으며, 전체 자료에서 2회 이상 등장한 표현만으로 감정표현 후보들을 추출하였다. 이러한 규칙 기반 필터링을 통해 최종 292개의 감정표현 사전들을 구축하였다. 최종 표현사전은 2-gram 120개, 3-gram 102개, 4-gram 54개, 5-gram 16개로 구성되었다.

본 절에서는 면담 문장 단위에서 감정표현이 함께 사용되는 양상을 확인하기 위해 공출현 기반 결합 분석을 수행하였다. 문장 내에서 동시에 등장한 감정표현 n-gram 쌍을 추출하고 각 표현 쌍에 대해 자카드 유사도와 PMI를 산출함으로써 표현 간 공출현의 빈도적 수준과 결합 강도를 확인하였다.

분석 결과, 자카드 유사도와 PMI 값이 높은 상위 표현 쌍은 동일한 어절이 후행 표현이나 조사, 서술어와 결합되며 확장된 형태로 나타났다. 예를 들어 ‘잠을 못’, ‘잠을 못 자는’, ‘아직도 해결’, ‘아직도 해결 안’ 등 동일한 의미 단위를 공유하며 문장 내에서 반복적으로 함께 사용되었다. 이러한 표현 쌍은 동일 문장 맥락에서 지속적으로 공통으로 출현하여 이에 따른 공출현 빈도와 자카드 유사도, PMI 값이 모두 높게 표 1로 산출되었다.

해당 결과는 면담 발화에서 감정표현이 발화의 흐름에 따라 점진적으로 확장되는 방식으로 사용된다는 점을 나타낸다. 감정표현은 조사나 후행 서술이 덧붙여지며 길어지고 구체화되며, 동일 어절을 중심으로 여러 확장 형태가 함께 사용되는 표현군이 나타난다. 이러한 표현 확장 양상은 면담 텍스트 전반에서 반복적으로 관찰된다는 점을 정량적으로 확인하였다.

종합하면 공출현 기반 결합 분석에서 도출된 상위 표현 쌍은 감정 표현이 문장 맥락에 따라 확장되며 사용되는 언어적 양상이 반복되는 것을 보여준다. 이는 본 연구의 n-gram 기반 접근이 재난 이후 경험이 언어적으로 서술되는 과정에서 어떻게 나타나는지를 파악하였다.

그림 1은 공출현 기반 감정표현 쌍의 자카드 유사도와 PMI 분포를 나타낸 것이다. 상위 표현 쌍은 자카드 유사도와 PMI가 동시에 높은 영역에 집중되어 있다. 자카드 유사도는 0~1 사이의 값이 나타나며, 값이 높은 경우 해당 쌍이 동일 문장 맥락에서 함께 사용되는 빈도가 높음을 보여준다. PMI 값은 0보다 작은 경우 같이 나오는 경향이 낮음을 표현하고, 0~1인 경우 약한 결합 강도를 보이며 값이 1 이상인 경우 높은 결합을 보이는 것으로 해석한다.

추가적으로 n-gram 기반 감정표현 추출 과정에서 동일 어절을 공유하는 확장 표현이 다수 생성되는 특성을 고려하여, 문장 단위 공출현 분석의 한계를 보완하고자 분석 단위를 단계적으로 확장하였다.

턴(turn) 단위 분석은 동일 발화 흐름 내에서 연속적으로 언급되는 감정표현 간의 결합을 포착하기 위한 목적으로 표 2

표 1. 문장 기반 공출현 결과 (부분집합 포함)

Table 1. Sentence-level co-occurrence results (including subset relations)

Phrase A	Phrase B	Co-occur	Jacard	PMI
2차 가해 그런	가해 그런	4	1	5.15618
못 자는	잠을 못 자는	3	1	5.44386
화가 나는	화가 나는 거예요	3	1	5.44386
공주 트라우마	트라우마 센터	4	0.8	4.93303
못 자니까	잠을 못 자니까	4	0.8	4.93303
...	...	...	...	...
못 자고	잠을 못	3	0.14286	2.53514
아예 잠을 못	잠을 못	3	0.13636	3.45143
...	...	...	...	...

\*The above figure presents original Korean text excerpts to preserve the linguistic characteristics of the source data.

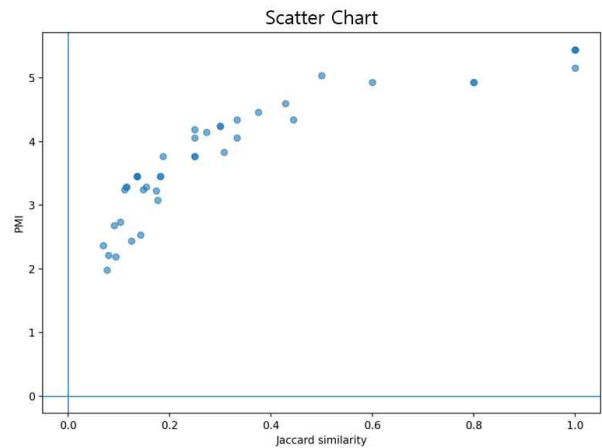


그림 1. 문장 기반 표현의 산점도 결과

Fig. 1. Scatter plot of sentence-level expressions

표 2. 발화 기반 공출현 결과(부분집합 제외)

Table 2. Utterance-level co-occurrence results (excluding subset relations)

Phrase A	Phrase B	Co-occur	Jacard	PMI
참사 이후에	피해자 분들	5	0.45455	3.20422
...	...	...	...	...
너무 힘들어	안 되겠다	2	0.4	3.81036
...	...	...	...	...
2차 가해	피해자 분들	3	0.21429	2.60639
...	...	...	...	...
못 자고 그러니까	잠을 못	2	0.13333	2.89407
...	...	...	...	...

\*The above figure presents original Korean text excerpts to preserve the linguistic characteristics of the source data.

표 3. 윈도우 사이즈(k=2) 기반 공출현 결과(부분집합 제외)

Table 3. Co-occurrence results based on window size (k=2) (excluding subset relations)

Phrase A	Phrase B	Co-occur	Jacard	PMI
못 자고 그러니까	잠을 못 잤어요	2	0.5	5.03551
...	...	...	...	...
계속 물이	생각을 계속했던 것	3	0.33333	4.17246
...	...	...	...	...
갑자기 이제	되게 힘들어하세요	2	0.25	4.05468
...	...	...	...	...
못 자고	잠을 못 잤어요	2	0.13333	3.49506
...	...	...	...	...
너무 힘들어서	잠도 못 자고	2	0.10526	3.01323
...	...	...	...	...
공주 트라우마	참사 이후에	2	0.03125	0.43815

\*The above figure presents original Korean text excerpts to preserve the linguistic characteristics of the source data.

로 수행하고, 윈도우 사이즈 기반 분석은 문장 경계를 넘어 근접한 맥락에서 함께 사용되는 표현 간의 연결성을 확인하기 위해 표 3으로 적용되었다. 이를 통해 동일 문장 내 동시 등장 여부에 의해 제한되던 공출현 분석의 범위를 확장하고자 하였다. 또한, n-gram 표현이 다른 n-gram 표현의 부분집합이 되는 경우도 제외하여 추가적인 분석을 수행하였다.

그러나 분석 단위를 문장, 턴, 윈도우로 확장하여 적용한 결과에서도 상위 결합으로 도출된 표현 쌍의 양상은 유사하게 나타났다. 즉, 자카드 유사도와 PMI 값이 높은 표현 쌍의 다수는 서로 다른 감정표현 간의 결합이라기 보다는, 동일한 어절을 중심으로 조사나 서술어가 결합되며 확장된 형태의 표현들이 반복적으로 나타났다.

이러한 결과는 공출현 분석 단위의 확장만으로는 감정표현 간의 느슨한 의미적 연관 구조를 충분히 포착하기 어렵다는 점을 시사하며, 본 연구에서 관찰된 표현 결합 양상이 발화 맥락의 확장보다는 표현 추출 단위의 특성과 밀접하였다.

자주 추출된 표현들은 수면, 장소 및 사건 맥락과 관련된 표현군으로 감정 상태를 직접 지시하는 단일 감정어보다는 경험의 지속성이나 상황적 맥락을 서술하는 표현이 사용되는 양상을 보였다. 이러한 표현들은 동일 문장 또는 인접한 발화 맥락에서 반복적으로 등장함으로써 높은 공출현 빈도와 결합 강도를 나타냈고 면담 발화에서 감정이 개별 단어 단위로 분석되기보다 서술적 맥락에서 나타났다.

공출현 기반 감정표현 결합 분석 결과, 상위 표현들은 서로 다른 감정 간의 결합이라기보다 동일한 어절을 중심으로 조사와 서술어가 결합되며 확장된 형태로 반복적으로 나타나는

양상을 보였다. 또한 이러한 양상은 문장, 발화, 윈도우 단위로 분석 범위를 확장한 경우에도 유사하게 유지되었다. 이는 감정표현이 개별 감정어로 분리되어 나타나기보다 특정 상태를 중심으로 반복적으로 서술되는 경향이 있음을 보여준다.

결과적으로 이러한 특징 표현들은 재난 피해자의 발화에서는 현재 상태를 나타내는 표현, 사건 경험과 관련된 표현, 그리고 트라우마 센터와 같은 지원 체계와 관련된 표현이 함께 나타나는 경향이 확인되었다. 특히 수면 문제와 같은 현재 상태 표현과 사건 이후를 지시하는 표현 그리고 지원 경험을 나타내는 표현이 반복적으로 결합되고 사용되는 양상이 관찰되었다.

4-2 감정표현 사용의 빈도 및 다양성

감정표현의 사용 양상은 개인 및 집단 단위에서 뚜렷한 차이를 나타내었다. 표 4는 집단별 감정표현 사용량의 기초 통계를 산출한 결과이다. 유가족 집단의 경우 평균 감정표현 사용량은 39.67회, 생존자 집단은 36.00회로 나타났다. 두 집단의 평균 사용 빈도는 유사한 수준으로 나타났으며, 생존자 집단의 표준편차는 18.63으로 유가족 집단의 11.31보다 크게 나타났다. 또한 생존자 집단의 감정표현 사용 빈도는 최소 14회에서 최대 69회까지 분포하였으며, 유가족 집단의 경우 최소 21회에서 최대 55회로 나타났다. 이러한 결과는 재난 이후 경험이 동일한 집단 내에서도 빈도의 개인차가 존재함을 나타낸다.

표 4. 그룹별 감정표현의 통계 요약

Table 4. Summary statistics of emotion-expression usage by group

Group	N	Mean hit	Standard Deviation	Min	Max
Bereaved	12	39.67	11.31	21	55
Survivor	9	36.00	18.63	14	69

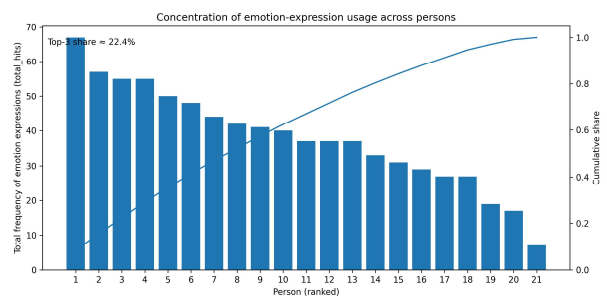


그림 2. 개인별 감정표현 빈도

Fig. 2. Concentration of emotion-expression usage across persons

그림 2는 개인별 감정표현 사용 빈도를 내림차순으로 정렬한 결과이다. 최종 구축된 292개의 감정표현 사전을 기준으로 전체 면담 발화를 분석한 결과, 감정표현은 중복 사용을

포함하여 총 800회 매칭되었다. 이중 상위 3명의 참여자는 179회의 감정표현을 사용하였고 전체의 22.38%를 차지하였다. 상위 5명의 참여자는 총 284회를 사용하였으며 전체의 35.5%를 차지하였다. 반면 다수의 참여자는 상대적으로 제한된 빈도로 감정표현을 사용하는 양상을 보였다. 이러한 결과는 감정표현 사용 빈도가 개인 간에 균등하게 분포되지 않으며, 일부 참여자에게 집중되는 경향이 있음을 나타낸다.

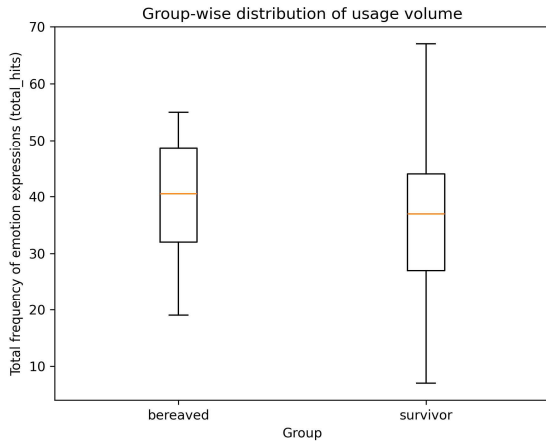


그림 3. 그룹별(유가족, 생존자) 감정표현 사용 분포  
Fig. 3. Group-wise distribution of usage volume

그림 3은 개인 단위에서 관찰된 감정표현 사용 빈도의 분포 특성을 집단 단위로 구분하여 나타낸 결과이다. 분석 결과 유가족 집단은 감정표현 사용 빈도의 중앙값이 다소 높게 나타났으며, 비교적 안정적인 분포를 보였다. 반면 생존자 집단은 집단 내 분산이 크게 나타나 개인 간 편차가 두드러지는 양상을 보였다. 이러한 결과는 감정표현 사용 빈도가 집단 간에서도 큰 차이를 보이며, 특히 생존자 집단에서 두드러지는 것을 나타낸다.

표 5. 개인별 표현의 다양성

Table 5. Individual-level expression diversity statistics

Metric	Value
N	21
Mean entropy	3.212
Standard deviation	0.418
Minimum entropy	1.946
Maximum entropy	3.809

감정표현 사용 빈도와 함께 표현의 다양성을 확인하기 위해 개인별 엔트로피 값을 산출하였다. 엔트로피 값은 특정 표현에 집중되지 않고 다양한 표현이 얼마나 사용되었는지를 나타내는 지표로 활용된다. 표 5는 평균 엔트로피 값과 표준편차, 최대값과 최소값을 나타낸다. 평균 엔트로피 값은 3.212로 나타났으며, 최소 1.946에서 최대 3.809까지 분포하였다. 또한 표준편차는 0.418로 나타나 참여자 간 표현 사용의 다양성에서도 차이가 나타났다. 엔트로피 값이 낮은 경

우에는 제한된 표현이 반복적으로 사용되는 경향이 나타났으며, 높은 경우는 자신의 경험을 다양한 표현을 사용하여 서술하는 것을 의미한다. 이러한 결과는 동일한 재난 경험 이후에도 감정표현을 구성하고 사용하는 방식이 개인별로 다르게 나타남을 보여준다.

표 6. 그룹별 표현의 다양성

Table 6. Group-level expression diversity statistics

Group	N	Mean Entropy	Standard deviation	Minimum entropy	Maximum entropy
Bereaved	12	3.290	0.301	2.801	3.809
Survivor	9	3.110	0.543	1.946	3.721

집단별 감정표현 사용의 다양성을 비교하기 위해 그룹 단위 엔트로피 값을 확인하였다. 표 6은 분석 결과를 나타낸다. 유가족 집단의 경우 평균 엔트로피는 3.290이며, 생존자 집단은 3.110으로 나타났다. 이는 유가족 집단에서 감정표현이 상대적으로 더 다양한 형태로 분산되어 있음을 나타낸다. 반면 생존자 집단은 평균 엔트로피가 다소 낮게 나타나는 양상을 보였다. 또한 생존자 집단의 표준편차는 0.543으로 유가족 집단의 0.301보다 크게 나타나 집단 내 개인 간의 표현 다양성 차이가 크게 나타났다. 엔트로피 범위에서도 유가족 집단은 비교적 안정적이었으나 생존자 그룹은 최솟값과 최댓값의 차이가 크게 나타나는 것을 확인하였다. 이러한 결과는 동일 그룹 내에서도 표현 양상의 차이가 큰 것을 보여준다.

이러한 결과를 종합하면 감정표현의 사용은 개인 및 집단에 따라 서로 다른 분포 양상으로 나타났다. 감정표현의 사용량은 개인별 편차가 크게 나타났으며, 집단 내에서도 분포 범위의 차이가 관찰되었다. 특히, 생존자 집단은 사용 빈도의 범위가 넓게 나타나 개인 간 표현 사용량의 차이가 상대적으로 크다는 것을 확인할 수 있다. 반면 감정표현의 다양성 분석 결과에서는 유가족 집단이 평균적으로 더 다양한 표현을 사용하는 경향을 보였으며, 생존자 집단은 개인 간 표현 구성 방식의 편차가 크게 나타났다. 이러한 차이는 재난 피해자의 발화가 경험을 서술하는 과정에서 빈도와 구성 방식 모두에서 서로 다른 양상으로 나타날 수 있음을 보여준다.

#### 4-3 감정표현 사용 패턴 및 유사성

감정표현의 사용은 인터뷰 대상자 간에 전반적으로 낮은 유사성을 보였다. 각 대상자의 면담 텍스트에서 감정표현 분포 벡터를 구성하고 코사인 유사도를 산출한 결과, 다수의 대상자 쌍은 낮은 유사도 구간에 분포하였다. 이는 감정표현 사용이 집단적으로 수렴된 하나의 전형적인 패턴으로 형성되기 보다는 개별 대상자의 경험과 서술 맥락에 상이함을 의미한다. 일부 대상자 쌍에서는 상대적으로 높은 유사도가 관찰되었으나, 전체적인 경향을 대표하기보다는 공통성이 존재하는 사례로 해석된다. 그림 4는 21명의 인터뷰 대상자 각각을 서로 2명씩 묶은 쌍의 유사도 분포이다. 전체 쌍의 다수는 0.0

에서 0.15 사이에 분포하였고, 평균 유사도는 0.095로 나타났다.

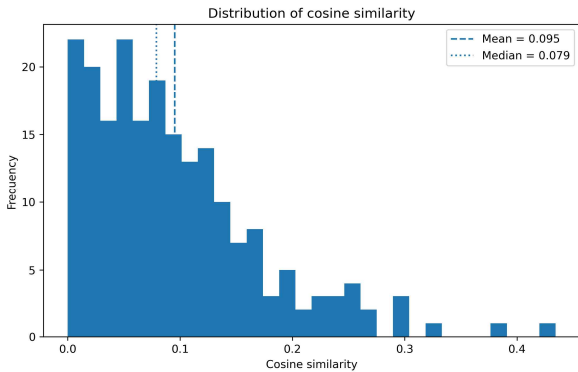


그림 4. 코사인 유사도 분포  
Fig. 4. Cosine similarity distribution

감정표현의 결합 양상을 확인하기 위해 문장 단위 공출현 관계로 표현 네트워크를 구성하고 커뮤니티 탐지를 수행하였다. 각 문장에서 함께 등장한 표현 쌍의 공출현 빈도를 집계한 뒤 자카드 유사도를 산출하였으며, 이를 바탕으로 일정 기준 이상인 표현 쌍만 연결하는 유사도 네트워크를 구축하였다. 그러나 자카드 유사도 분포는 대부분이 0이며, 극단적으로 최소화해 나타나 전역 임계값(threshold=0.08) 기반으로 네트워크를 구성하는 경우 전체 표현 쌍 중 해당 임계값을 초과하는 경우는 약 0.5% 수준에 불과하였다. 이는 유사도 분포가 극단적으로 최소화하고 일부 강한 결합 관계만이 네트워크를 이루는 것으로 나타났다.

이에 각 표현이 상대적으로 높은 유사도를 보이는 이웃 노드를 유지하는 top-k 전략을 적용하여 국소적 연결성을 보존한 네트워크를 구축하였다. 이는 표현 간 강한 결합과 함께 주변 맥락적 연관성까지 반영하여 전역 기반 임계값 기준 집합보다 풍부한 표현 집합을 확인하였다. 이후 탐욕적 모듈러리티 기반 커뮤니티 탐지를 수행하여 표현 군집을 그림 5와 같이 도출하였다.

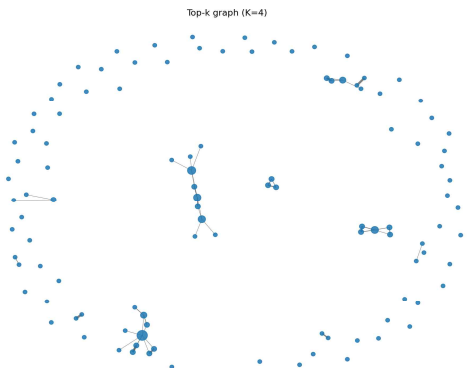


그림 5. 커뮤니티 군집 표현 그래프(k=4)  
Fig. 5. Community clustering graph (k=4)

k 값은 2~6 범위에서 민감도 분석을 수행하였다. 표 7은 민감도 분석 결과이다. 분석 결과 엣지 수와 최대 커뮤니티 규모는 큰 변동없이 안정적으로 유지되었다. 엣지의 수는 41~43 범위에서 유지되었고, 최대 커뮤니티 규모는 10으로 동일하게 나타났다. 또한, 모듈러리티 값 또한 0.85 수준에서 큰 변동을 보이지 않았다.

표 7. k 값 별 네트워크 지표 요약

Table 7. Summary of network metrics across k values

K	Edge	Largest community	Modularity(Q)
2	41	10	0.8606
3	42	10	0.8508
4	43	10	0.8509
5	43	10	0.8509
6	43	10	0.8509

이에 따라, k=4로 설정하여 커뮤니티를 도출하였다. 도출된 네트워크는 총 81개의 커뮤니티로 분할되었고, 표 8은 이중 상위 커뮤니티는 각각 10개, 8개, 5개의 표현으로 구성된 것을 나타낸다. 상위 커뮤니티를 구성하는 표현을 살펴보면, 감정표현은 단일한 형태가 아니라 특정한 상태와 경험을 중심으로 결합된 서술 단위가 나타났다. 특히, “그 이후에”, “그 참사 이후에” 등 사건 이후 시간적 지속성에 대한 표현, “잠도 못 자고”, “아예 잠을 못” 등 현재 상태와 관련된 어려움을 직접적으로 드러내는 표현이 관찰되었다. 또한 “계속”과 같은 반복 표현과 “트라우마 센터” 등과 같은 표현에서 개인이 겪는 어려움이 지속되고 있으며, 외부 지원 체계를 나타내는 표현이 등장한다. 이러한 결과는 재난 피해자의 발화가 사건 자체의 단순한 설명이 아니라 사건 이후 현재 상태, 반복되는 경험으로 인해 어려움을 호소하고 있는 것으로 확인되었다. 이러한 감정표현은 특정 상태와 경험을 중심으로 조직된 서술 단위로 나타나는 특성을 보이며, 이를 통해 재난 이후 경험이 어떻게 언어적 표현으로 발화되는지를 확인하였다.

표 8. 상위 커뮤니티 요약(K=4)

Table 8. Summary of top communities (k=4)

Size	Representative expressions (selected example)
10	그 이후에, 그 이후에 이제, 그 참사 이후에, ...
8	잠도 못 자고, 못 자는, 아예 잠을 못, 이제 잠을 못, ...
5	공주 트라우마 센터, 트라우마 센터에서, ...
3	계속 물이, 계속 이제, 이제 계속

\*The above figure presents original Korean text excerpts to preserve the linguistic characteristics of the source data.

## V. 한계점 및 향후 연구

본 연구는 재난 피해자의 면담 텍스트를 기반으로 감정표현의 분포와 결합 양상을 분석하고, 이를 통해 표현이 나타나는 방식과 의미를 함께 해석하고자 하였다. 이러한 접근은 감정표현을 단순히 분류하거나 빈도로 파악하는 것에서 나아가

재난 이후 경험이 언어적으로 어떻게 서술되는지를 확인하였다는 의의가 있다.

그러나 본 연구에서 제시한 해석은 면담 발화에 나타난 표현을 기반으로 한 것으로 개별 참여자의 심리 상태나 회복 수준을 직접적으로 판단하는 데에는 한계가 있다. 특히 현재 상태를 나타내는 표현, 사건 경험과의 결합, 표현의 반복과 같은 양상은 사건 이후에도 재난 경험을 지속적으로 서술하고 있음을 보여주는 단서로 해석될 수 있으나, 특정한 심리적 상태로 단정하기에는 제한이 존재한다. 또한, n-gram 기반 공출현 분석은 동일 어절을 중심으로 확장된 표현이 반복적으로 도출되는 특성이 있으며, 이로 인해서 서로 다른 감정이나 경험 간의 연관 구조를 포착하는 것에 한계가 있다.

더불어 재난 사건의 면담자료는 표본의 규모와 특정 재난에 국한된 결과로 일반화에는 신중한 해석이 필요하다. 추후 연구에서는 다양한 재난 유형과 확대된 자료를 활용하여 다른 자료에서도 일관된 맥락을 검토해야 한다. 나아가 향후 연구에서는 공출현 기반 분석을 넘어 임베딩 기반 접근이나 대규모 언어모델을 활용하여 표현 간 의미적 관계를 보다 확장된 맥락에서 분석하여 현재 상태, 사건 경험, 지원 체계와 같은 표현 간의 관계를 정밀히 파악할 필요가 있다.

## VI. 결 론

본 연구는 재난 피해자의 면담 텍스트를 기반으로 감정표현의 분포와 결합 양상을 분석하고, 이를 통해 피해자의 경험이 실제 발화에서 어떠한 형태로 나타나는지를 확인하였다. 특히 사전 정의된 감정어에 의존하지 않고 면담 과정에서 생성된 표현을 분석 대상으로 설정함으로써, 재난 이후 경험이 언어적으로 어떻게 구성되고 서술되는지를 실증적으로 제시하였다.

분석 결과 감정표현은 개별 감정어로 분리되어 나타나기보다, “그 이후에”, “그 참사 이후에”, “잠도 못 자고”와 같이 현재 상태와 사건 이후의 경험이 결합된 형태로 반복적으로 나타나는 양상이 확인되었다. 또한 “계속”과 같은 반복 표현과 “트라우마 센터”와 같은 지원 체계 관련 표현이 함께 나타난다는 점은 재난 이후 경험이 일회적 사건이 아니라 현재의 상태와 연결되어 지속적으로 서술되고 있음을 보여준다.

또한, 감정표현의 사용 양상은 개인 및 집단 간의 차이를 보였으며, 유가족 집단은 평균적으로 보다 다양한 표현을 사용하였고, 생존자 집단은 개인 간 표현 사용량과 구성 방식에서 편차가 크게 나타났다. 이러한 차이는 재난 이후 경험이 단일한 방식으로 서술되는 것이 아니라, 개인 및 집단에 따라 서로 다른 방식으로 구성되고 표현될 수 있음을 시사한다.

이러한 결과를 종합하면 재난 피해자는 사건 이후에도 현재 상태와 연결된 형태로 지속적인 어려움을 호소하고 있으며, 수면 문제와 반복적 회상과 같은 증상 관련 표현이 실제 발화에서 반복적으로 확인된다. 이는 재난 이후 회복 과정에

서 나타나는 증상 호소를 고려한 심리적 지원이 필요함을 보여주는 결과이다. 따라서 재난 이후 지원은 단기적 대응에 그치지 않고, 장기적으로 피해자의 상태를 반영한 심리적 치료 및 지원이 지속적으로 이루어질 수 있도록 국가적 차원의 제도적 기반이 마련되어야 한다.

## 감사의 글

본 연구는 행정안전부 국립재난안전연구원의 지원(“재난 피해 회복지원 시스템 개발 기획연구”, “NDMI-기본-2026-10-01”)에 의해 수행되었습니다. 이에 감사드립니다.

## 참고문헌

- [1] A. Matsukawa, S. Nagamatsu, R. Ohtsuka, and H. Hayashi, “Disaster Resilience Scale for Individuals: A Fundamental Requirement for a Disaster-Resilient Society,” *International Journal of Disaster Risk Reduction*, Vol. 107, 104405, 2024.
- [2] J. Kim, H. Lim, D. Nam, J. Sim, S. Lee, H. Kim, and S. Park, “Enhancing Temporary Housing Models for Disaster Resilience: Insights Drawn from Post-Disaster Experiences in Korea,” *Sustainability*, Vol. 17, No. 20, 9225, 2025.
- [3] L. Zou, N. S. N. Lam, H. Cai, and Y. Qiang, “Mining Twitter Data for Improved Understanding of Disaster Resilience,” *Annals of the American Association of Geographers*, Vol. 108, No. 5, pp. 1422-1441, 2018.
- [4] M. Moghadas, A. Fekete, A. Rajabifard, and T. Kötter, “The Wisdom of Crowds for Improved Disaster Resilience: A Near-Real-Time Analysis of Crowdsourced Social Media Data on the 2021 Flood in Germany,” *GeoJournal*, Vol. 88, pp. 4215-4241, 2023.
- [5] C. S. Holling, “Resilience and Stability of Ecological Systems,” *Annual Review of Ecology and Systematics*, Vol. 4, pp. 1-23, 1973.
- [6] M. Linnenluecke and A. Griffiths, “Beyond Adaptation: Resilience for Business in Light of Climate Change and Weather Extremes,” *Business & Society*, Vol. 49, No. 3, pp. 477-511, 2010.
- [7] B. Rouhanizadeh, S. Kermanshachi, and T. J. Nipa, “Exploratory Analysis of Barriers to Effective Post-Disaster Recovery,” *International Journal of Disaster Risk Reduction*, Vol. 50, 101735, 2020.
- [8] J.-M. Park and S.-M. Bae, “Impact of Depressive, Anxiety, and PTSD Symptoms in Disaster Victims on Quality of Life: The Moderating Effect of Perceived Community Resilience,” *International Journal of Disaster Risk Reduction*, Vol. 69, 102749, 2022.
- [9] E. Macleod, T. Heffernan, L.-M. Greenwood, L. Walker, J.

Lane, S. K. Stanley, ... and S. Sutherland, "Predictors of Individual Mental Health and Psychological Resilience after Australia's 2019-2020 Bushfires," *Australian & New Zealand Journal of Psychiatry*, Vol. 58, No. 1, pp. 58-69, 2024.

- [10] A. Suryani and S. Soedarso, "Language of Disaster and Society: The Socio-Emotional Functions of Language after Natural Disaster," in *Proceedings of the 5th Prasasti International Seminar on Linguistics (PRASASTI 2019)*, pp. 34-40, 2019.
- [11] S. K. Kang, H. Yu, and Y. J. Lee, "Analyzing Disaster Response Terminologies by Text Mining and Social Network Analysis," *Information Systems Review*, Vol. 18, No. 1, pp. 141-155, 2016.
- [12] S. Z. Razavi and M. Rahbari, "Understanding Reactions to Natural Disasters: A Text Mining Approach to Analyze Social Media Content," in *Proceedings of 2020 7th International Conference on Social Networks Analysis, Management and Security (SNAMS)*, Paris, France, pp. 1-7, 2020.
- [13] Z. Zhao, T. Liu, S. Li, B. Li, and X. Du, "Ngram2vec: Learning Improved Word Representations from Ngram Co-Occurrence Statistics," in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, Copenhagen, Denmark, pp. 244-253, 2017.
- [14] L. H. Lin, S. B. Miles, and N. A. Smith, "Natural Language Processing for Analyzing Disaster Recovery Trends Expressed in Large Text Corpora," in *Proceedings of the 2018 IEEE Global Humanitarian Technology Conference (GHTC)*, San Jose: CA, pp. 1-8, 2018.
- [15] M. C. Rahmadan, A. N. Hidayanto, D. S. Ekasari, and B. Purwandari, "Sentiment Analysis and Topic Modelling Using the Lda Method Related to the Flood Disaster in Jakarta on Twitter," in *Proceedings of the 2020 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*, Jakarta, Indonesia, pp. 126-130, 2020.
- [16] S. Deb and A. K. Chanda, "Comparative Analysis of Contextual and Context-Free Embeddings in Disaster Prediction from Twitter Data," *Machine Learning with Applications*, Vol. 7, 100253, 2022.
- [17] M. Bayer, "Information Overload in Crisis Management: Bilingual Evaluation of Embedding Models for Clustering Social Media Posts in Emergencies," in *Deep Learning in Textual Low-Data Regimes for Cybersecurity*, Wiesbaden, Springer Fachmedien Wiesbaden, pp. 67-87, 2025.
- [18] A. Martini and D.-M. Buda, "Analysing Affects and Emotions in Tourist E-Mail Interviews: A Case in

Post-Disaster Tohoku, Japan," *Current Issues in Tourism*, Vol. 22, No. 19, pp. 2353-2364, 2019.



**김지호(Ji-Ho Kim)**

2022년 : 중앙대학교 대학원 (공학석사)

2022년~현 재: 국립재난안전연구원

2023년~2025년: 부산대학교 정보융합공학과 컴퓨터공학전공 박사수료

※ 관심분야 : GIS(Geographic Information System), 그래프 이론(Graph Theory), 텍스트 마이닝(Text mining), 재난대피(Disaster Evacuation), 알고리즘(Algorithm) 등



**이소희(So-Hee Lee)**

2009년 : 일본 츠크바대학 대학원 (공학박사)

2014년~현 재: 국립재난안전연구원

2009년~2013년: 도쿄대학 공간정보과학연구센터/대학원 특임 연구원

※ 관심분야 : 도시 방재 및 회복력(Urban Disaster Risk Reduction and Resilience), 재난구조·복구 및 피해자 지원(Disaster Relief, Recovery, and Victim Support), 공간정보과학(Spatial Information Science)



**임혜선(Hye-Sun Lim)**

2016년 : 광운대학교 대학원 (심리학석사)

2016년~현 재: 국립재난안전연구원

※ 관심분야 : 재난피해자(Disaster Victim), 재난회복력(Disaster Resilience), 면담조사(Interview Survey) 등



**남동엽(Dong-Yup Nam)**

2015년 : 광운대학교 대학원  
(심리학 석사)

2018년~현 재: 국립재난안전연구원

※ 관심분야 : 재난피해자(Diaster Victim), 재난구호(Disaster Relief), 심리지원(Psychological Support) 등



**황우석(Woo-Suk Hwang)**

2011년 : 서울과학기술대학교 학사  
(안경광학 전공)

2014년 : 국가평생교육진흥원 행정학사  
(사회복지학 전공)

2018년 : 한양대학교 대학원  
보건학석사 (직업 및  
환경보건 전공)

2018년~현 재: 국립재난안전연구원 선임연구원

※ 관심분야 : 재난안전(Disaster Safety), 재난피해 평가  
(Disaster Damage Assessment), 회복수준  
(Resilience Level), 재난예경보시스템(Early  
Warning and Alert Systems) 등



**양수연(Su-Yeon Yang)**

2025년 : 동의대학교 대학원  
(재난관리학석사)

2025년~현 재: 국립재난안전연구원

※ 관심분야 : 재난관리(Disaster Management), 재난구호  
(Disaster Relief), 안전취약계층(Safety  
Vulnerable Groups) 등



**박상현(Sang-Hyun Park)**

2018년 : 국립서울과학기술대학교 안전  
공학과 (공학박사)

2003년 : 국립서울과학기술대학교 안전  
공학과 (공학석사)

2006년~현 재: 국립재난안전연구원 공업연구원

2024년~현 재: 국립재난안전연구원 재난회복연구센터 센터장

※ 관심분야 : 재난구호(Disaster Relief), 재난대피(Disaster Evacuation), 재난회복력(Disaster Resilience) 등