



HAL
open science

**”Cheese!”: A corpus of face-to-face French interactions.
A case study for analyzing smiling and conversational
humor**

Béatrice Priego-Valverde, Brigitte Bigi, Mary Amoyal

► **To cite this version:**

Béatrice Priego-Valverde, Brigitte Bigi, Mary Amoyal. ”Cheese!”: A corpus of face-to-face French interactions. A case study for analyzing smiling and conversational humor. Language Resources and Evaluation Conference, 2020, Marseille, France. pp.460-468. hal-02565645

HAL Id: hal-02565645

<https://hal.science/hal-02565645v1>

Submitted on 13 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

"Cheese!": a corpus of face-to-face French interactions. A case study for analyzing smiling and conversational humor

Béatrice Priego-Valverde, Brigitte Bigi, Mary Amoyal

Aix-Marseille Univ, CNRS, LPL

5 avenue Pasteur, Aix-en-Provence, France

{beatrice.priego-valverde;brigitte.bigi;mary.amoyal}@univ-amu.fr

Abstract

The aim of this article is to present Cheese! a conversational corpus and, as an illustration, an example of study that such a multimodal corpus allows. Cheese! consists of 11 French face-to-face conversations lasting around 15 minutes each. In this article, the methodology used to collect and enrich the corpus "Cheese!" is detailed: experimental protocol, technical choices, transcription, semi-automatic annotations, manual annotations of smiling and humor. Then, an exploratory study investigating the links between smiling and humor is proposed. Based on the analysis of two interactions of "Cheese!", two questions are asked: (1) Does smiling frame humor? (2) Does smiling have an impact on its success or failure? This study does not claim to fully answer these questions. Rather, it exemplifies which kinds of investigation are made possible by such a corpus. Although the experimental design of Cheese! has been elaborated specifically for the study of smiling and humor in conversations, the creation of high quality data set and the methodology presented can be replicated and applied to the analysis of many other conversational activities (such as narration or explanation) and other multimodal phenomena.

Keywords: Smiling, humor, conversation, annotation

1. Introduction

Cheese! is an audio-video conversational corpus recorded in 2016 at the LPL - Laboratoire Parole et Langage, Aix-en-Provence, France, at the CEP - Centre for Speech Experimentation¹. It consists of 11 mixed and non-mixed dyadic interactions, lasting around 15 minutes each. Cheese! was first collected in order to make a cross-cultural comparison of smiling during humorous productions between American English and French (Priego-Valverde et al., 2018). Consequently, it was recorded following the American protocol, as closely as possible, especially concerning the tasks given to the participants (see section 2.2).

The aim of this article is twofold. First, it presents the technical choices and the experimental protocol used to construct the Cheese! corpus. So far, 5 interactions of the corpus have been semi-manually annotated using different tools: the corpus has been automatically segmented into Inter-Pausal units using SPPAS (Bigi, 2015) and manually transcribed using SPPAS and then Praat (Boersma and Weenink, 2018). Several automatic annotations were then generated using SPPAS and MarsaTag (Rauzy et al., 2014) software tools. Among these 5 interactions, 2 of them have been manually annotated: participants' smiles using ELAN (Brugman et al., 2004), and humorous sequences using Praat. Second, we propose an exploratory study investigating the links between smiling and humor. Based on the analysis of two interactions, two questions will be asked: (1) Does smiling frame humor? (2) Does smiling have an impact on its success or failure?

2. Description of the corpus

2.1. Participants

The 22 participants of the corpus were students in Linguistics at Aix-Marseille University. The two participants of each interaction were in the same class and were also friends outside the university. All were French native students, from 20 to 40 years old. All were informed of the purpose of the data collection when it was entirely finished and all signed a written consent form (available on Ortolang) before the recordings.

2.2. Experimental protocol

Two tasks were given to the participants. First, they were asked to read each other a canned joke chosen by the researchers (see annex). Second, they were asked to speak as freely as they wished until the end of the interaction. These two tasks were given in order to compare the sequences of canned jokes and conversational humor both in French and American English corpora (Priego-Valverde et al., 2018). But in this article, only the sequences of conversational humor were analyzed.

The participants were recorded in a soundproof room where they were seated face-to-face as illustrated in Figure 1. Participants were fitted with two headset microphones (AKG-C520) optimally positioned so as not to hide the mouth. These microphones were connected by XLR to the RME Fireface UC, which is connected with a USB cable to a PC using Audacity software. The recordings are all sampled at 48,000 Hz in PCM signed 16 bits big-endian files (.wav). Two cameras were positioned in such a way that each participant was shown from the front. A video editing program was used to merge the two videos into a single one (Figure 2) and to embed the high-quality sound of the microphones. All the videos are 1920x1080 px sampled at 25 fps, with H264 codec embedded in a mp4 container.

All these primary data are stored in Ortolang, the Open

¹ CEP is a shared experimental platform for the collection and analysis of data for the study of speech production and perception. Among others, the CEP allows to gather a wide range of audio and video data in its anechoic room.

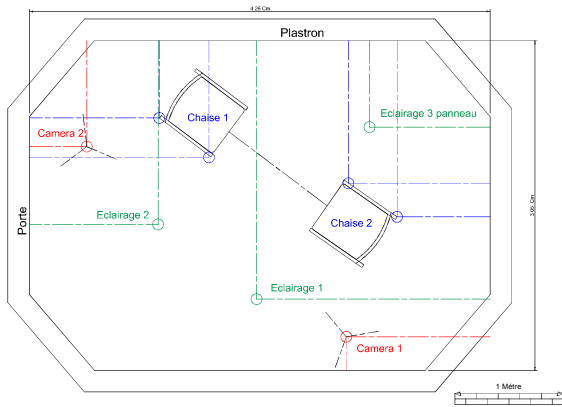


Figure 1: Design of the room and participants



Figure 2: Experimental design of Cheese!

Resources and TOols for LANGuage repository (Priego-Valverde, Béatrice, 2016). A demo of 46 seconds is publicly available and the whole corpus can be shared on-demand, after registration. However, audio files re-sampled at 8,000 Hz and video files reduced to 363x206 px are publicly available and can be freely downloaded without registration.

3. Annotation of the data

So far, 5 interactions have been enriched with annotations. From the audio file of each speaker, the Inter-Pausal Units (IPUs) were automatically extracted. Then, the orthographic transcription was done manually by one of the authors and both the IPUs and the orthographic transcription were verified by another one of the authors. Several automatic annotations were then generated, such as time-aligned tokens, phonemes, syllables, TGA, morpho-syntax, syntactic categories, and lemmas. Other annotations were also provided manually from the audio signals or from the videos, including smiles and humor.

3.1. Inter-Pausal Units

Inter-Pausal Units are the result of a sounding/silence segmentation. They are widely used for large corpora in order to facilitate speech alignments and for the analyses of speech-like prosody (Peshkov et al., 2012).

The “Search for IPUs” automatic annotation of SPPAS software tool was used on the 10 audio files of 5 dialogues. The following parameters were used:

- Minimum silence duration is 200 ms, a common value for French,
- Minimum IPU duration is 100 ms which is appropriate to properly find the isolated feedback like “mh”,
- Shift-left the beginning of the IPUs is 20 ms allows to avoid truncating the first word,
- Shift-right the end of the IPUs is 20ms to avoid truncating the final word.

The resulting annotations were manually verified with Praat. It resulted in 10 files with the expected IPUs and 10 files with the IPUs automatically found by SPPAS. The automatic system has been satisfactorily assessed on its ability to meet the expected result (Bigi and Priego-Valverde, 2019).

3.2. Orthographic transcription

The orthographic transcription was performed manually by one of the authors. Only the IPUs were listened with “IPUScriber” tool of SPPAS. The transcription was verified by another one of the authors with Praat.

In spontaneous speech, numerous phenomena occur such as hesitations, repetitions, non-standard elisions, reduction phenomena, truncated words, and more generally, non-standard pronunciations. Events like laughter, noises and filled pauses also occur very frequently in spontaneous speech (Bigi and Meunier, 2018). The transcription of Cheese! includes all these, and for the annotation of these, we relied on the Enriched Orthographic Transcription convention (EOT) of SPPAS software tool:

- a noise is annotated ‘*’; it can be a breath, a cough or an unintelligible segment, ...
- laughter is annotated by a ‘@’
- a short pause is annotated by a ‘+’
- a broken word is annotated with a ‘-’ at the end of the token string
- an elision is mentioned between parenthesis, like thi(s)
- an unexpected pronunciation is annotated with brackets like this [example, eczap]
- an unexpected liaison is surrounded by ‘=’
- a comment of the transcriber is annotated with braces like {this is a comment}

Two transcriptions can then be automatically derived from the EOT (Bigi et al., 2012). The standard one contains the standard form of words and the faked one contains a phonetic transcription of words. For example, the phrase « thi(s) [example, eczap] “ has the standard form “this example” and its faked form is “thi eczap”. The first one is human-readable and relevant for syntactic or lexical analyses but the second one provides a better grapheme-to-phoneme conversion and is therefore more temporally aligned with the audio stream.

3.3. Automatic annotations

From the manual transcriptions within the IPUs and from the audio signals, we generated some of the automatic annotations available in SPPAS, named as Text normalization (Bigi, 2014), Phonetization (Bigi, 2016), Alignment (Bigi

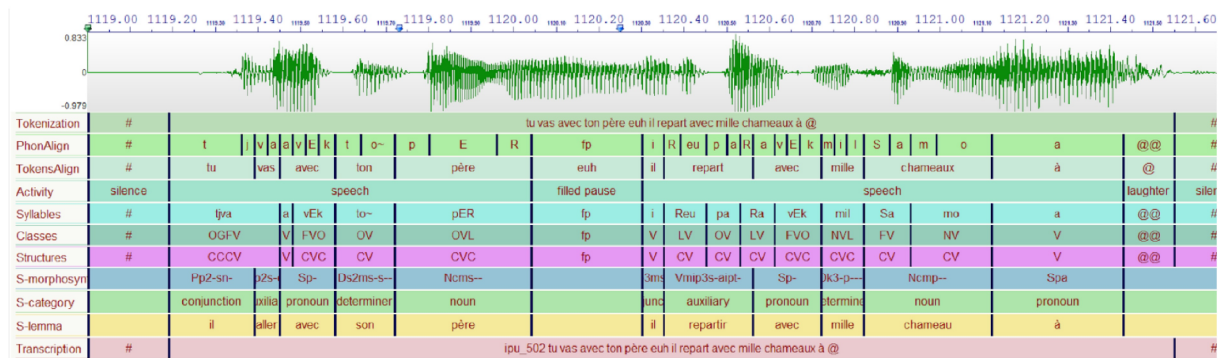


Figure 3: Example of some of the automatic annotations generated by SPPAS and MarsaTag software tools from the Transcription (last tier) and the audio file.

and Meunier, 2018), Syllabification (Bigi et al., 2010) and Activity. Each of them resulted in a time-aligned set of annotation tiers. One of the specificities of such system is that the filled-pause is represented by a specific acoustic model; and the same for the laughter and the noise. It results in accurate segmentation of spontaneous speech, as shown in (Bigi and Meunier, 2018).

Figure 3 displays several of such tiers: only the last one with the EOT was manually constructed. All the other tiers were automatically generated: the Tokenization tier is one of tiers resulting of the text normalization process; the PhonAlign tier represents time-aligned sounds (phonemes, laughter, filled pauses, short pauses and noises); the TokensAlign tier contains time-aligned tokens like words, truncated words, filled pauses, etc; the Activity tier is the main activities such as speech or silence; Syllables/Classes/Structures are the result of the syllabification process. The part-of-speech (POS) tags were automatically identified using MarsaTag (Rauzy et al., 2014). It considers 9 parts-of-speech categories: adjective, adverb, auxiliary, conjunction, determiner, noun, preposition, pronoun, and verb. The result is represented by the tiers S-morphosyntax, S-category and S-lemma.

3.4. Manual annotations of smiles

The smiles that occurred during two interactions, (MA_PC and JS_CL), were manually annotated, using the “Smiling Intensity Scale” (SIS) (Gironzetti et al., 2016). The SIS progressively measures the smile intensity from 0 (neutral face) to 4 (laughter), based on Action Units (AUs) detailed by the Facial Action Coding System (FACS) (Ekman and Friesen, 1978). Table 1 represents the 5 levels of smile intensity from the corpus.

In applying this scale, manual annotations of smiles were performed with ELAN software on each participant. Instead of deciding where a smile starts and ends, each interaction was divided into 400 ms intervals. This sampling rate has been chosen as this is considered the time necessary to produce or perceive a complex gesture such as smiling (Sanders and Sanders, 2013; Heerey and Crossley, 2013). Then, each interval was assigned a smile intensity. This predefined interval at which a smile category must be assigned allows to allow for reduced subjectivity in annotating the beginning or end of a smile. Furthermore, it is easier to perform counter-coding on intervals of the same duration and

location. 2,610 smile intensities were annotated in MA_PC and 2,475 in JS_CL. This annotation protocol allows us to precisely analyze the evolution of each participant’s smile. A double coding was carried out on both interactions to validate the reliability of these annotations and the relative objectivity of the scale used. We then calculated Cohen’s Kappa (Cohen, 1960), a statistical measure used to compare the annotations of two judges. Both inter-annotator agreement rates were qualified as excellent: 0.87 for MA_PC and 0.89 for JS_CL.

3.5. Manual annotations of humor

Humorous instances of the same two interactions (MA_PC and JS_CL) were manually annotated by one of the authors using Praat (Figure 4). First, the entire humorous sequences were annotated and classified either as canned joke (CJ) or conversational humor (CH). Then, within each sequence of conversational humor, humorous items were annotated. Two categories of criteria were applied to recognize humor. The first one stems from the General Theory of Verbal Humor (Attardo and Raskin, 1991; Attardo, 2001), such as the “script opposition”, and the “logical mechanism” (Attardo, 2001). The second one stems from previous studies highlighting linguistic devices frequently used by humor in conversation such as punning, pinning, repetition (see section 1.3) and references to the participants’ common ground (Priego-Valverde, 2003).

Finally, the analysis of each reaction to a humorous item allows us to distribute each of them in failed (Fa) and successful humor (Su). Three negative reactions were observed (Priego-Valverde, 2020): humor acknowledged but answered seriously (A), humor ignored (I), and humor explicitly rejected (R). When none of these reactions were observed, the humorous items were qualified as successful.

4. Lexical coverage

From the result of the automatic text normalization, we evaluated that the 5 dialogues of Cheese! are made up of 2,130 different words representing 20,201 occurrences. In addition to these words, the dialogues include the following items:

- 3,196 silences;
- 334 pauses;
- 176 noises;

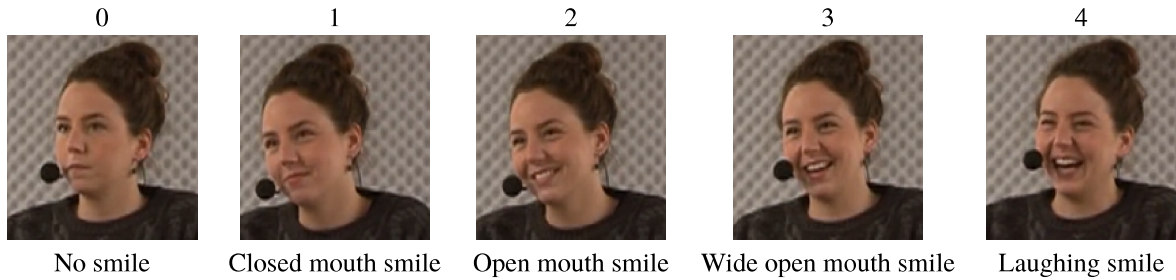


Table 1: Smiling Intensity Scale (Gironzetti et al., 2016). The illustrations of each level are pictures extracted from the Cheese! corpus

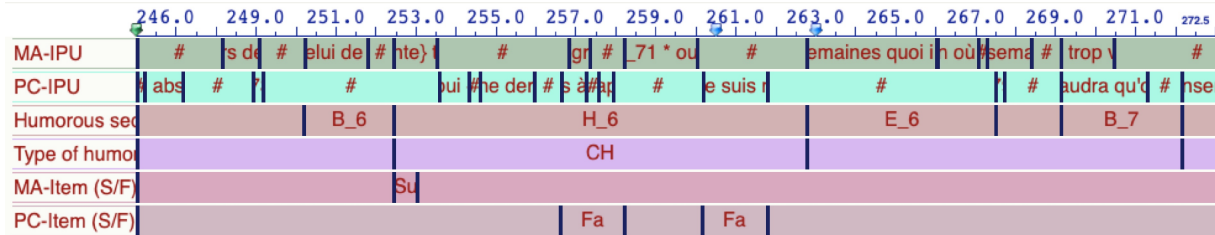


Figure 4: Example of some of the manual annotations

- 401 instances of laughter;
- 598 filled pauses ("euh");
- 133 different truncated words occurring 265 times together.

Of the 2,130 words of the vocabulary, 1,027 (48.2%) occur only once. The 10 most frequent are:

- 663 est (to be at present third person singular)
- 610 c' (it)
- 581 tu (you, singular informal)
- 569 ouais (yeah)
- 516 je (I)
- 426 pas (not)
- 426 et (and)
- 382 de (of)
- 329 ça (that/this)
- 315 la (the, feminine gender)

These 10 most frequent words occur 4,817 times, so they cover 23.8% of the vocabulary. Figure 5 illustrates this relation among the words and their occurrences. We can see that a vocabulary size of 40 words covers about 50% of all spoken words.

5. Exploratory analysis of smiling and humor in Cheese!

5.1. Smiling in conversation

Lots of information is emitted through gestures, facial expressions, postures, and other modalities (Birdwhistell and Lacoste, 1968; Graham and Argyle, 1975). Out of several modalities available and used by the interlocutors, smiling is “the most frequent conversational facial expression” (Cosnier and Kerbrat-Orecchioni, 1987). Even if smiling

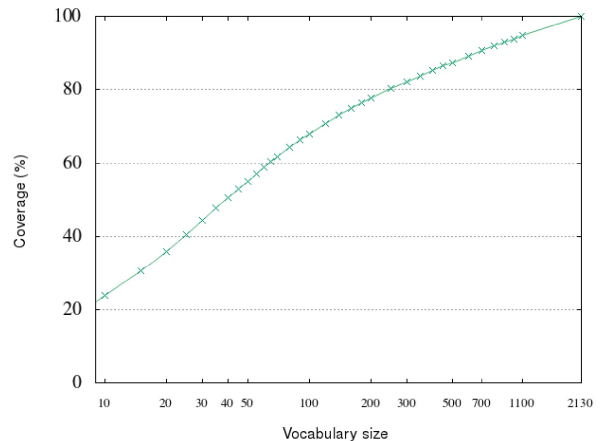


Figure 5: Lexical coverage

is often studied for its emotional functions such as joy (Ekman, 1984) or happiness (Fernández-Dols and Ruiz-Belda, 1995), smiling has also been studied for its interactive functions such as feedback expressions (Jensen, 2015). As smiling can also convey pragmatic and interactive functions it is also considered as a “facial gesture” (Bavelas et al., 2014). That is why smile is considered here a communicative gesture, in the perspective of analyzing the relationship between smiling and humor in conversation. Smile has been mostly analyzed in a binary way as it has been described depending on its authenticity with the dichotomy “social” vs. “authentic” or even “Duchenne” and “Non-duchenne” smiling (Ekman and Friesen, 1975). Furthermore, smiling has been mostly analyzed in terms of presence/absence. To report the complexity of this facial expression, smiling will be investigated here in terms of its presence but also in terms of intensities. While the presence of smiling will be investigated as a potential marker of humor, its various intensities allows us to deeper analyze the role of smiling during humor

and the impact of smiling in the success or failure of humor. In that aim, our approach of smiling relies on the Smiling Intensity Scale (SIS) (Gironzetti et al., 2016), an annotation system based on the Action Units from the FACS (Ekman and Friesen, 1978). This scale divides smiling into 5 degrees of intensities, from neutral face to laughing smile.

5.2. (Failed) humor in conversation

Since Norrick's seminal work on "conversational joking" (Norrick, 1993) studies on humor in conversation are still increasing. Many of them focus on the social / relational functions of conversational humor, seen as a mixed phenomenon going "from bonding to biting" (Boxer and Cortés-Conde, 1997). Another kind of work focuses on the humorous devices used in conversations, such as punning (Norrick, 1993), pinning (Priego-Valverde, 2016), and repetitions (Tannen, 1989). Although far less frequent, studies on failed humor began to really appear in the 2000s. The most important part of such studies focuses mainly on the hearer's negative reactions (among others (Hay, 1994; Hay, 2001; Eisterhold et al., 2006; Attardo, 2002; Bell, 2009)), or on the hearer's failure to support the speaker's humor. For example, according to (Bell, 2015) who applies Hay's model of humor support (2001) to failed humor, i.e. humor not supported, the 5 reasons that may lead to a failure of humor are a lack of recognition, understanding, appreciation, agreement and/or engagement from the hearer. More recently, focusing on the humor production and not perception, Priego-Valverde (Priego-Valverde, 2019; Priego-Valverde, 2020) analyzed failed humor produced by the participant who is in the position of hearer. The way that conversational constraints weight on on humor production is highlighted.

5.3. Laughter and smiling as multimodal markers of humor in conversation

Links between laughter and humor have been regularly mentioned since the seventies in Conversation Analysis. Considered a marker of humor (Sacks, 1974) laughter was studied by (Jefferson, 1979) as a device in conversation used by the speaker to show his/her humorous intention. In line with such studies, laughter has even been considered "the contextualization cue for humor par excellence" (Kotthoff, 2000), and its absence has been seen as a mark of failure of humor (Norrick, 1993). However, the relationship between laughter and humor is more than questionable. Firstly, humor does not necessarily trigger laughter and laughter is not always provoked by humor (Attardo, 1994; Foot, 2017; Morreall, 2001; Priego-Valverde, 2003). Secondly, it has been shown that not only does the lack of laughter not necessarily mean that humor has failed, but it can also be seen as a support strategy (Hay, 2001).

More recently, studies on smiling and humor have begun to appear. The most significant have been conducted in computer-mediated conversations. In (Attardo et al., 2013), for example, smiling synchronicity was studied in relation with the presence or absence of humor, hypothesizing on the function of "framing" of smiling. This hypothesis was deepened by (Gironzetti et al., 2016) adding the role of the synchronicity of smiling between participants.

This "framing" role of the synchronicity of smiling has also been highlighted in French face-to-face conversations (Priego-Valverde, 2017; Priego-Valverde and Bigi, 2016). In (Gironzetti et al., 2019), the role of smiling and smiling synchronicity, as a discourse marker has been also investigated. They show that, even if smiling is present in the whole conversation, its intensity is higher during humor productions. Thus, the authors hypothesize that an increase in smiling intensity is a more significant way to frame an exchange as humorous than the sole presence of smiling.

In line with these previous studies, the present exploratory study investigates the links between smiling and humor. Two questions will be asked: (1) Is smiling a marker of humor? (2) Does smiling diminish the risk that humor will fail? In order to provide potential answers to these questions, methodological decisions have been made:

- Smiling was isolated from other facial behaviors such as neutral face and laughter, using the SIS;
- When a humorous item included at least one occurrence of smiling, and even associated with other facial behavior, it was analyzed as smiling;
- Humorous productions were analyzed in two ways: Firstly, in order to present a general picture of the two interactions analyzed, the total duration of the humorous productions in the entirety of the interactions was calculated based on the humorous sequences extracted. These sequences include the humorous items produced by the speakers and the hearers' reactions. In each interaction, the two first humorous sequences were excluded because they correspond to the canned joke that the participants had to read to each other, and not to conversational humor as freely appearing in any conversation. Secondly, for the rest of the analysis focusing on humorous productions (and not reactions to these), only the humorous items included in each humorous sequence were studied.

5.4. General overview of the data

5.4.1. Duration of the humorous productions

Based on this methodology, the total duration of humorous productions was calculated comparing the duration of the humorous sequences with the duration of the non-humorous sequences: 2.6 min in MA_PC (for a total interaction lasting 17.4 min) and 7.03 min in JS_CL (for a total interaction lasting 16.5 min). Such a high difference between the two interactions can be explained by two factors: a gender effect (MA_PC is a mixed dyad while JS_CL are two women), and the conversational topics developed (in MA_PC, they spend a long time talking about their future exams).

5.4.2. Frequency of the humorous items

The frequency of the humorous items produced by each participant was then calculated (Table 2).

Here again, high inter-individual differences appear, especially between MA and CL. MA is the lesser prolific participant. He is also the only man and the participant who, because he is a good student, advises PC on the way she has to study. Thus, the two same factors than showed above could explain this difference.

Participants	Nb. of humorous items	Avg.
MA	23	1.3/min
PC	32	1.8/min
JS	39	2.36/min
CL	53	3.2/min

Table 2: Number of humorous items by participant

5.5. Relationship between smiling and humor

As smiling is frequent in conversation, its sole presence cannot be an indicator *per se* of humor. We compare humorous and non-humorous sequences, hypothesizing that smiling intensity will be higher in the former, in line with (Gironzetti et al., 2019).

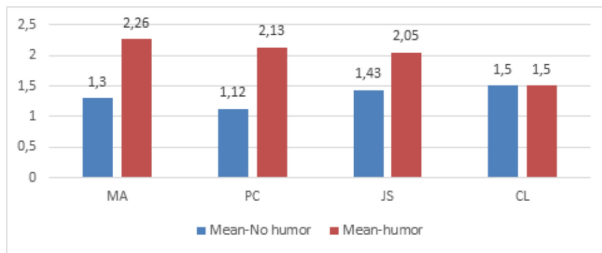


Figure 6: Mean smiling intensity during humorous and non-humorous segments

Figure 6 shows that mean smiling intensity is mainly higher during humorous sequences, which is consistent with (Gironzetti et al., 2019). However, such result must be interpreted with caution because one participant (CL) displayed the exact same mean smiling intensity in both humorous and non-humorous sequences. If a higher mean smiling intensity is more frequent during humorous sequences, it is not systematic.

Then, in order to investigate the role of smiling as a marker of humor, smiling was isolated from other facial behaviors (neutral face and laughter) displayed by each participant while producing their humorous items.

Figure 7 shows that smiling is the most frequent facial behavior displayed by all the participants while producing humor. Laughter is far less frequent and, unsurprisingly, displaying a neutral face is rare. Thus, observing smiling in a binary way (presence/absence), tends to confirm the assumption that smiling is a marker of humor.

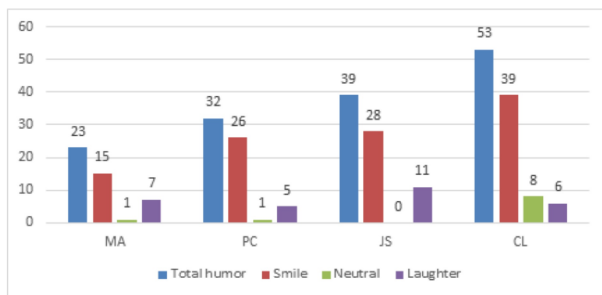


Figure 7: Number of facial behaviors per participant during humor production

5.6. Potential impact of smiling on the success or failure of humor

Considering that these results tend to indicate that both intensity mean (figure 6) and presence/absence of smiling (figure 7) are markers of humor, the impact of smiling on the success or failure of humor was then investigated. In other words, does smiling reduce the risk that humor will fail? To answer this question, we compared the number of failed and successful humorous items produced by all the participants.

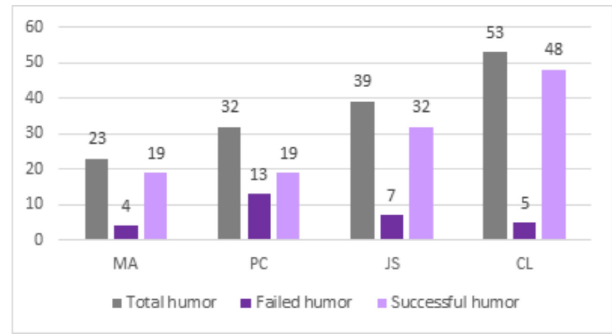


Figure 8: Number of failed and successful humorous items by participant

Figure 8 shows that failed humor is less frequent than successful humor, which is consistent with other studies conducted on another corpus (Priego-Valverde, 2020). This tendency is very clear, except for one participant (PC) whose results are more balanced.

Then, the percentage of smiles present in failed and successful humorous items was also compared (Figures 9 and 10). They were isolated from the two other facial behaviors displayed by participants, i.e. neutral face and laughter.

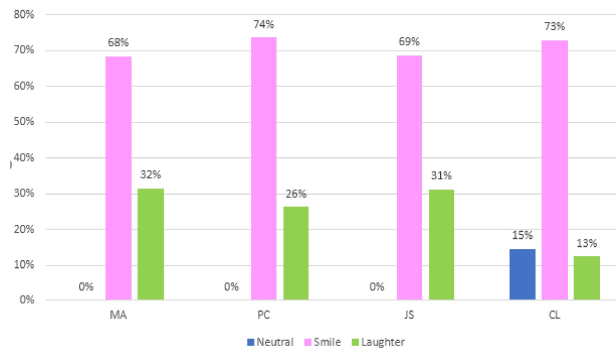


Figure 9: Percentage of successful humorous items produced with and without smile

Figure 9 shows that all the participants display many more smiles than other facial behaviors during successful humor, and more than laughter. Unsurprisingly, the neutral face is almost nonexistent. At first glance, such a result seems to confirm that smiling favors the success of humor. However, Figure 10 mitigates these results. It shows that all the participants display also many more smiles than other facial behaviors during failed humor. This results therefore, tend to confirm the high rate of smiling during humorous sequences and, consequently, its role as a marker of humor.

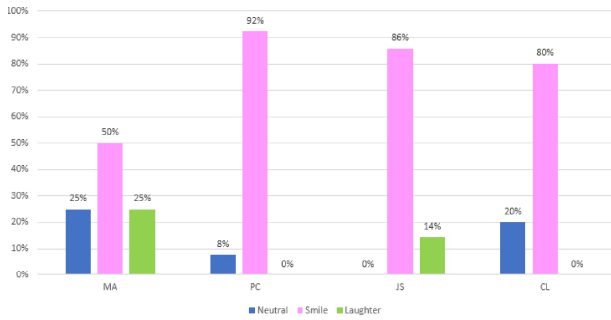


Figure 10: Percentage of failed humorous items produced with and without smile

But they do not show that smiling reduces the risk for humor to fail.

Finally, the mean smiling intensity displayed both in failed and successful humorous items has been calculated.



Figure 11: Mean smiling intensity displayed during failed (Fa) and successful (Su) humor

Figure 11 shows that three participants display a higher mean smiling intensity during successful humor. However, MA's one is lower during successful humor than failed humor. MA being the only male participant, this difference could be explained by a gender effect.

When comparing figures 9 and 11, the results show that, with successful humor, most of the participants display more smiles than other non-verbal behaviors with also a higher mean intensity than with failed humor, except for MA.

When comparing figures 10 and 11, the results show that, with failed humor, most of the participants display also more smiles than other facial behaviors, but with a lower mean intensity than with successful humor, except for MA.

Such results show that mean smiling intensity seems to be a more robust criterium than the sole presence or absence of smile to evaluate the impact of smiling on the success or failure of humor. However, based on only four participants including one exception (MA), such a result cannot be more precise.

When comparing Figure 6 and 11, the results show that all the participants whether they produce failed or successful humor, display a higher mean smiling intensity during humorous items than during non-humorous sequences. This result tends to show that mean smiling intensity works also as a marker of humor.

Finally, CL's very low mean smiling intensity with failed humor diminishes a lot her mean smiling intensity during humor. It could explain why she displays the same smiling intensity whether or not she produces humor (see Figure 6).

6. Discussion

The methodology used to annotate both failed and successful humor has already been used on another corpus (Priego-Valverde, 2020). It is thus replicable. The methodology used for annotating smiling through its different degrees of intensities has also been used in previous studies on humor in American English (Attardo et al., 2013; Gironzetti et al., 2016; Gironzetti et al., 2019), on humor in French (Priego-Valverde and Bigi, 2016; Priego-Valverde, 2017), and on studies about topic transitions in French conversations (Amoyal and Priego-Valverde, 2019). More recently, this methodology made it possible to conduct research on automatic detection of smiles (Rauzy & Amoyal, submit). Moreover, investigating smiling behavior with the "Smiling Intensity Scale" (Gironzetti et al., 2016) describing smile according to 5 levels of intensities is particularly relevant not only to distinguish smiling from other non-verbal behaviors such as the neutral face and laughter, but also to go beyond the binary analysis of smiling in terms of presence vs. absence.

This study provides some clues about the role of smiling as a maker of humor. First, smiling is the most frequent facial behavior used in humor. It is used more than neutral face and more than laughter. Considering that the links between laughter and humor have been much more thoroughly investigated, the role of smiling in humor could have been underestimated by previous studies. The second result concerns the mean smiling intensity which is higher in humorous sequences than in non-humorous sequences. This result is consistent with previous studies (see section 1.3). In other words, both the presence and the mean smiling intensity seem to be a marker of humor.

Focusing on whether or not smiling diminishes the risk for humor to fail, the results are less meaningful. On the one hand, smiling is the most present facial behavior both in successful and in failed humor. Thus, it does not seem to reduce the risk of failure. On the other hand, mean smiling intensity is higher in successful humor than in failed humor. This result could indicate that mean smiling intensity has a larger impact of the success or failure of humor than the simple presence of smiling. However, such a result cannot be considered meaningful either because, this only holds for three participants. The fourth participant actually displayed a higher mean smiling intensity during failed humor. One exception among only four participants represents too large of an uncertainty to draw any clear conclusion about the impact of mean smiling intensity on the success or failure of humor.

7. Conclusion

This paper presents the Cheese! corpus consisting of a set of 11 face-to-face interactions. Five of these interactions are already largely annotated with enriched orthographic transcription, time-aligned phonemes, syllables, words, part-of-speech, etc. A first observation concerns the restricted

vocabulary used: 40 words cover 50% of the dialogs. So far, manual annotations of smiling and humor have been provided for two interactions (MA_PC and JS_CL).

The results of this exploratory study need to be interpreted with caution, and further research is needed to deepen our understanding of the interplay between smiling and humor. Thus, when all the interactions are transcribed and annotated, further studies will take two directions. First, the same analysis will be carried out on the entire corpus in order to confirm (or not) the role of smiling and mean smiling intensity as a marker of humor. The real impact of mean smiling intensity on the success or failure of humor will be also investigated more in depth. Second, certain factors which may explain some of the variations in our results will be explored, such as gender and conversational topics developed by the participants.

The experimental protocol used for Cheese! has led to the creation of a high-quality data set that opens many perspectives for future research. The separation of the audio files facilitates their transcriptions, allowing for phonetic and prosody analyses. The conversational setting allows for the analysis of conversational phenomena. Finally, the quality of the videos allows for future analysis of smiling, and more generally, of other non-verbal phenomena such as hand gestures, nods, and postures.

8. Acknowledgements

The authors wish to thank the CEP - Centre d'Expérimentation de la Parole, the shared experimental platform for its help in creating the recordings.

9. Annexes

Frog joke

An engineer was crossing a road one day when a frog called out to him and said, "If you kiss me, I'll turn into a beautiful princess."

He bent over, picked up the frog and put it in his pocket. The frog spoke up again and said, "If you kiss me and turn me back into a beautiful princess, I will stay with you for one week."

The engineer took the frog out of his pocket, smiled at it and returned it to the pocket. The frog then cried out, "If you kiss me and turn me back into a princess, I'll stay with you and do ANYTHING you want."

Again the engineer took the frog out, smiled at it and put it back into his pocket. Finally, the frog asked, "What is the matter? I've told you I'm a beautiful princess that I'll stay with you for a week and do anything you want. Why won't you kiss me?"

The engineer said, "Look I'm an engineer. I don't have time for a girlfriend, but a talking frog, now that's cool."

Donkey joke

A car was involved in an accident in a street. As expected a large crowd gathered. A newspaper reporter, anxious to get his story could not get near the car.

Being a clever sort, he started shouting loudly, "Let me through! Let me through! I am the son of the victim."

The crowd made way for him.

Lying in front of the car was a donkey.

10. Bibliographical References

- Amoyal, M. and Priego-Valverde, B. (2019). Smiling for negotiating topic transitions in french conversation. In *Gesture and Speech in Interaction*, Paderborn, Germany.
- Attardo, S. and Raskin, V. (1991). Script theory revis (it) ed: Joke similarity and joke representation model. *Humor-International Journal of Humor Research*, 4(3-4):293–348.
- Attardo, S., Pickering, L., Lomotey, F., and Menjo, S. (2013). Multimodality in conversational humor. *Review of Cognitive Linguistics. Published under the auspices of the Spanish Cognitive Linguistics Association*, 11(2):402–416.
- Attardo, S. (1994). Linguistic theories of humor (vol. 1). *Berlin: De Gruyter Mouton. Barsoux, JL (1996). Why organisations need humour. European Management Journal*, 14(5):500–508.
- Attardo, S. (2001). *Humorous Texts: A Semantic and Pragmatic Analysis*. Humor research. Bod Third Party Titles.
- Attardo, S. (2002). Humor, Irony and their Communication: from mode adoption to failure of detection. *Say not to Say: New perspectives on miscommunication*, 3:159–179.
- Bavelas, J., Gerwing, J., and Healing, S. (2014). *Hand and facial gestures in conversational interaction*. Academic Press.
- Bell, N. D. (2009). Responses to failed humor. *Journal of Pragmatics*, 41(9):1825–1836.
- Bell, N. (2015). *We are not amused: Failed humor in interaction*, volume 10. Walter de Gruyter GmbH & Co KG.
- Bigi, B. and Meunier, C. (2018). Automatic segmentation of spontaneous speech. *Revista de Estudos da Linguagem. International Thematic Issue: Speech Segmentation*, 26(4).
- Bigi, B. and Priego-Valverde, B. (2019). Search for interpausal units: application to cheese! corpus. In *9th Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics*, pages 289–293, Poznań, Poland.
- Bigi, B., Meunier, C., Nesterenko, I., and Bertrand, R. (2010). Automatic detection of syllable boundaries in spontaneous speech. In *7th International conference on Language Resources and Evaluation*, pages 3285–3292, La Valette, Malta.
- Bigi, B., Péri, P., and Bertrand, R. (2012). Orthographic transcription: which enrichment is required for phonetization? In *Proceedings of the Eight International Conference on Language Resources and Evaluation*, pages 1756–1763, Istanbul, Turkey. European Language Resources Association (ELRA).
- Bigi, B. (2014). A multilingual text normalization approach. *Human Language Technology Challenges for Computer Science and Linguistics*, LNAI-8387:515–526.
- Bigi, B. (2015). SPPAS - Multi-lingual Approaches to the Automatic Annotation of Speech. *The Phonetician*, 111–112:54–69.
- Bigi, B. (2016). A phonetization approach for the forced-alignment task in SPPAS. *Human Language Technology*.

- Challenges for Computer Science and Linguistics*, LNAI-9561:397–410.
- Birdwhistell, R. L. and Lacoste, M. (1968). L'analyse kinésique. *Langages*, 3(10):101–106.
- Boersma, P. and Weenink, D. (2018). Praat: doing phonetics by computer [computer program], version 6.0.37, retrieved 14 march 2018 from <http://www.praat.org/>.
- Boxer, D. and Cortés-Conde, F. (1997). From bonding to biting: Conversational joking and identity display. *Journal of Pragmatics*, 27(3):275–294.
- Brugman, H., Russel, A., and Nijmegen, X. (2004). Annotating multi-media/multi-modal resources with elan. In *LREC*.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46.
- Cosnier, J. and Kerbrat-Orecchioni, C. (1987). *Décrire la conversation avec R. Bouchard, L. Fontaney, M.-M. de Gaulmyn, S. Rémi-Giraud, Ch. Rittaut-Hutinet*, volume 15. Presses universitaires de Lyon.
- Eisterhold, J., Attardo, S., and Boxer, D. (2006). Reactions to irony in discourse: evidence for the least disruption principle. *Journal of Pragmatics*, 38(8):1239–1256.
- Ekman, P. and Friesen, W. (1975). *Unmasking the face : a guide to recognizing emotions from facial clues*. Englewood Cliffs : Prentice-hall.
- Ekman, P. and Friesen, W. V. (1978). *Facial Action Coding System: Manual*, volume 1-2. Consulting Psychologists Press.
- Ekman, P. (1984). Expression and the nature of emotion. *Approaches to emotion*, 3:19–344.
- Fernández-Dols, J.-M. and Ruiz-Belda, M.-A. (1995). Are smiles a sign of happiness? gold medal winners at the olympic games. *Journal of personality and social psychology*, 69(6):1113.
- Foot, H. (2017). *Humor and laughter: Theory, research and applications*. Routledge.
- Gironzetti, E., Attardo, S., and Pickering, L. (2016). Smiling, gaze, and humor in conversation. *Metapragmatics of humor: Current research trends*, 14:235.
- Gironzetti, E., Attardo, S., and Pickering, L. (2019). Smiling and the negotiation of humor in conversation. *Discourse Processes*, 56(7):496–512.
- Graham, J. A. and Argyle, M. (1975). A cross-cultural study of the communication of extra-verbal meaning by gestures (1). *International Journal of Psychology*, 10(1):57–67.
- Hay, J. (1994). Jocular abuse in mixed gender interaction. *Wellington Working Papers in Linguistics*, 6:26–55.
- Hay, J. (2001). *The pragmatics of humor support*. Walter de Gruyter.
- Heerey, E. A. and Crossley, H. M. (2013). Predictive and reactive mechanisms in smile reciprocity. *Psychological Science*, 24(8):1446–1455.
- Jefferson, G. (1979). A technique for inviting laughter and its subsequent acceptance/declination. *Everyday language: Studies in ethnomethodology*, pages 79–96.
- Jensen, M. (2015). Smile as feedback expressions in interpersonal interaction. *International Journal of Psychological Studies*, 7(4):95–105.
- Kotthoff, H. (2000). Gender and joking: On the complexities of women's image politics in humorous narratives. *Journal of pragmatics*, 32(1):55–80.
- Morreall, J. (2001). Sarcasm, irony, wordplay, and humor in the hebrew bible: A response to hershey friedman. *Humor*, 14(3):293–302.
- Norricks, N. R. (1993). *Conversational joking: Humor in everyday talk*. Indiana University Press.
- Peshkov, K., Prévot, L., Bertrand, R., Rauzy, S., Blache, P., and Aix-En-Provence, F. (2012). Quantitative experiments on prosodic and discourse units in the corpus of interactional data. In *SemDial 2012 (SeineDial): The 16th Workshop on the Semantics and Pragmatics of Dialogue*, page 183.
- Priego-Valverde, B. and Bigi, B. (2016). Smiling behavior in humorous and non humorous conversations: a preliminary cross-cultural comparison between american english and french. In *International Society for Humor Studies Conference*, Dublin, Ireland.
- Priego-Valverde, Béatrice. (2016). Cheese! <https://hdl.handle.net/11403/cheese>.
- Priego-Valverde, B., Bigi, B., Attardo, S., Pickering, L., and Gironzetti, E. (2018). Is smiling during humor so obvious? A cross-cultural comparison of smiling behavior in humorous sequences in american english and french interactions. *Intercultural Pragmatics*, 15(4):563–591.
- Priego-Valverde, B. (2003). *L'humour dans la conversation familière: description et analyse linguistiques*. Editions L'Harmattan.
- Priego-Valverde, B. (2016). Teasing in casual conversations. *Metapragmatics of Humor: Current research trends*, 14:215.
- Priego-Valverde, B. (2017). Does smile can frame an utterance as humorous? An analysis of smiling behavior in conversations. In *iGesto'17*, Porto, Portugal.
- Priego-Valverde, B. (2019). The "dark side" of humor. In *Plenary at ISHS*, Austin.
- Priego-Valverde, B. (2020). 'Stop kidding, I'm serious': Failed humor in French conversations. *Script-based semantics foundations and applications. Essays in honor of Victor Raskin*.
- Rauzy, S., Montcheuil, G., and Blache, P. (2014). Marsatag, a tagger for french written texts and speech transcriptions. page 220, Second Asian Pacific Corpus linguistics Conference.
- Sacks, H. (1974). An analysis of the course of a joke's telling in conversation. *Explorations in the ethnography of speaking*.
- Sanders, A. F. and Sanders, A. (2013). *Elements of human performance: Reaction processes and attention in human skill*. Psychology Press.
- Tannen, D. (1989). Talking voices: Repetition. *Dialogue, and Imagery in Conversational*.