

Corpus Catalogue of Pauses and Hesitations: A Computational Analysis of Genderlect in Pakistani Talk Shows

¹ **Tahir Shah**, MPhil English (Linguistics), Department of English, Abdul Wali Khan University, Mardan. Email: tahir4ici@gmail.com

² **Dr. Mahrukh Shakir**, Assistant Professor, Department of English, Abdul Wali Khan University, Mardan. Email: mahrukh@awkum.edu.pk

³ **Dr. Abdul Waheed Qureshi**, Assistant Professor, Department of English, Abdul Wali Khan University, Mardan. Email: waheedqureshi@awkum.edu.pk

⁴ **Hifsa Syed**, MPhil English (Linguistics), Department of English, Abdul Wali Khan University, Mardan. Email: hifsasyed2016@gmail.com

⁵ **Waqas Ahmad**, MPhil English (Linguistics), Department of English, Abdul Wali Khan University, Mardan. Email: ahmadwaqas343@gmail.com

Date of Submission: 10th August 2021 Revised: 25th October 2021 Accepted: 13th December 2021

Abstract: Language is one of the unique characteristics of human beings which has been of immense interest to man throughout history. Linguistics, as a scientific study of human language, has now expanded and evolved into a wide range of areas after a great historical advancement. The research at hand aims to identify the similarities and differences in Pakistani genderlect based on the findings of the study. This is done by exploring the ways in which pauses and hesitations of Pakistani males are different from those of Pakistani females using English as a foreign language. Using conversation analysis approach of Sacks, Jefferson, and Schegloff (1974), and AntConc which is a freeware corpus analysis tool, the study of pauses and hesitations in the speech of Pakistani males and females was performed by employing the corpus of the purposively selected talk shows of Pakistan (Tahir Shah, 2021). The overall findings of this study indicate that the linguistic features i.e., pauses and hesitations in the conversation of Pakistani males and females are significantly different because of the different cultural style of English language as per genderlect theory of Deborah Tannen (1990).

Copyrights @Muk Publications

Keywords: *AntConc; Conversation Analysis; Corpus; Culture; Genderlect; Hesitations; Pauses*

INTRODUCTION

Background of the Study

A detailed investigation of native speakers' everyday speech reveals a wide range of phenomena, such as speech hesitation, protracted silences, nonverbal vocalizations, such as 'Aaa' and 'Mmm', repairs, and repetitions. Although, speakers employ them in conventional and unmarked ways that are compatible with native speaker hesitation patterns and norms, a large portion of this escapes, unnoticed by interlocutors. Even while speech from one or more interlocutors may be referred to as continuous, a minute of reflection will show that it actually is not. Speakers must at the very least pause their speaking to take a breath. In severe situations, their speech may become very erratic, with protracted pauses, unnecessary sounds, irrelevant phrases, or reformulations that slow down the transmission of a message. These kinds of discontinuities are a part of the hesitating phenomena. Speech production cannot be continuous because it is a muscular action; therefore, pauses and interruptions are necessary. Since, breathing is tied to phonation,

Vol. 13 No.2 December, 2021

International Journal of Computational Intelligence in Control

physiologically expected pauses occur during the inspiration phase of respiration. Pauses brought on by the desire to breathe again naturally are controlled by the speakers and occur more frequently during turn-taking and at the conclusion of intonation groups (Zellner, 1994; Viola & Madureira, 2008).

Men and women manage people differently, in the way that men give orders while women often soften their demands and remarks. It means that the use of language by men seems to emphasize the status. The way men speak differs from that of women in a sense that both are somehow influenced by their social construct that is gender which has an influence in one way or the other on language. Similarly, the speech of men contains dominating elements which are variously found in the use of language at different levels (Köhler, 2008).

Men aim to rise to the top by demonstrating their knowledge of the world. Women, on the other hand, are always trying to employ phrases that imply conversation and intimation. To sustain association and information communication, they employ “rapport talk” (p. 111). In comparison to men, they are more polite and exhibit unity, which has a long-lasting impression on other people’s thinking. Men’s interactions are structured in a hierarchical manner, with the goal of providing knowledge and thereby controlling the dialogue. Women frequently utilize such sentences as to show a power imbalance (Tannen, 1990).

Statement of the Research Problem

According to John Gray (1992), the reason why men and women fight so frequently is because they have forgotten that they are supposed to be different from one another. Therefore, the research at hand aims to explore the ways in which the pauses and hesitations of Pakistani males are different from Pakistani females, using English as a second language (Shah, 2021).

Objectives of the Study

The objectives of this research are given below.

- To explore the ways in which pauses and hesitations of Pakistani males differ from Pakistani females, using English as a second language.

- To identify the most frequently used filled pauses in Pakistani genderlect in talk shows.

Significance of the Study

This research is intended to reveal the frequencies of the use of pauses and hesitations in the English genderlect of Pakistanis that will enable them to know the causes of pauses and hesitations. It will also contribute to the already existing body of research available regarding language and gender. The findings of this study are supposed to give theoretical and practical contribution to the body of research in pauses and hesitations as representation of gender in the field of linguistics in the context of English Language. Besides, it will also enable people to see as to what extent speakers use pauses and hesitations of various types while speaking English as a second language in talk shows. In addition, findings of the study can be useful for Pakistani speakers of English as a genderlect; as they will come to know the use of English as genderlect.

Research Questions

- In what ways are pauses and hesitations of Pakistani males different from Pakistani females, using English as a second language?
- What are the most frequently used filled pauses in Pakistani genderlect in talk shows?

Literature Review

To establish the importance of pauses for the speaker, studies of filled and silent pauses over the previous two decades are reviewed. A brief historical introduction is followed by an analysis of the theoretical implications of pause location and a summary of the pertinent research. Moreover, elements related to cognition, emotional state, and social interaction are taken into account when evaluating the functional value of pauses.

First, pauses in spontaneous speech were examined from the standpoint of language creation with the notion that pauses (both filled and silent) represent issues with language elaboration (Rochester, 1973).

Some scholars (e.g. Goldman-Eisler, 1958) have suggested that hesitation pauses following the initial word of an utterance merely represent the transitional likelihood of each subsequent word. Others (e.g. Johnson, 1965; Boomer, 1965) claim that pauses represent both local and distant planning, such as planning by bigger stakeholders. The question of whether filled pauses and silent pauses represent various production processes was explored in several studies.

When pauses are long, Maclay and Osgood (1959) hypothesized, they are filled; however, Boomer's (1965) investigation refuted this claim. Levelt (1983) claimed that self-repair pauses and phrase editing reveal that speakers have limited control over the process of producing their speech and instead manage it through self-monitoring based on parsing one's own inner or overt speech.

A broad range of issues are covered by audience studies, including the reception process and tactics in various media formats and platforms; the fandom phenomenon; and the research of a group of people (Sullivan, 2013). There are several study initiatives that have used the general public as an analytic object in order to determine how viewers' profiles and perceptions of shows influence cultural hierarchies and social identification categories like gender, age, and race (Mittell, 2004).

According to Tannen (1990), men like to be the center of attention by demonstrating their erudition. While interacting with others, they are always attempting to climb the social hierarchy. While men are more likely to utilize conversational language, women are more likely to use phrases such as dialogue and intimation in discussion. For the sake of companionship and the exchange of information, they participate in rapport discourse. Women are more polite and united in their contacts with others than men, which leaves a lasting impression on their audience. Men use "report talk" (p. 111) while attempting to prove their rank. In order to assert supremacy over the discourse, their relationships are hierarchically arranged. It is usual for women to utilize terms such as power imbalance to demonstrate their lack of authority. In

women's conversation, the objective is to provide counsel as opposed to expressing her opinions openly. Men consider women's discourse to be useless because they perceive it to be empty rhetoric that does not provide any information, while males consistently seek to impart knowledge when speaking.

Khan et al. (2013) argue that no one can escape the influence of the media in modern culture since so many people like watching Pakistani plays on television. Tannen (1990) dispels the idea that males are superior to females in the public realm by analyzing the cultural disparities in how men and women are reared.

Methodology

Population and Sampling Procedure

Ten (10) complete talk shows in English language were purposively selected and downloaded from YouTube which are broadcasted by Pakistani channels. The hosts and guests in the talk shows were selected purposively to represent both the genders i.e., male and female, in order to find the gender differences of the hosts' and guests' linguistic features used in the talk shows.

The study is quantitative in nature that is, it used statistics to examine the data and the numbers produced were assured to produce an unbiased outcome that can be applied to a bigger population.

Data Collection

It is hoped that the results of the suggested quantitative research may be extrapolated to a broader population by using statistics to examine the data. Data was acquired from talk shows programs on Pakistani television networks as a sample for the research.

Table 1. Pakistani Talk Shows Taken from PTV World

S. No.	Talk Shows
1	Working Women in Different Fields
2	Sky is Limit
3	Annural Khalid Singer/App to Deliver Fresh Fruits & Vegetables
4	Women Hygiene Management/Innovations in Rehabilitation Technology & Research
5	WTM Student Education & Singing
6	Dialogue 8 May 2020
7	Learning Dynamics at Workplace after COVID-19 & Painting as a Passion
8	Perspective 05 06 2021
9	World Ocean Day/Singing
10	Blood Donation Awareness & Twin Sisters Blogger

Data Analysis Procedure

This research on the relationship between language and gender has been developed to analyze the frequencies of pauses and hesitations in the English genderlect of Pakistanis in talk shows using the classification of pauses and hesitations of Goldman-Eisler (1958). The conversation analysis approach of Sack, et al. (1974) was used in this study. The conversation analysis approach investigates the various functional devices employed in conversational exchanges and analyses talk in actions and interactions (Wilkinson & Kitzinger, 2017). The researchers used inductive data-driven analysis after transcription to look for reoccurring patterns of interaction. They identified regularities, rules, or models to describe these patterns based on the analysis according to the conversational analysis approach.

As the present study is a corpus based study, it is why the researchers used AntConc in order to avoid things unobserved/unnoticed. Corpus analysis is a type of text analysis that allows for large-scale comparisons of textual items. It permits

us to see things that the researcher(s) would not otherwise notice when reading the transcribed talk shows.

Data Analysis and Discussion

Gender Differences in Pauses and Hesitations

According to Goldman-Eisler’s (1958) linguistic category and speech execution classification, first, the data has been presented in tabular form according to psycho-linguistic category and speech execution and has been analyzed and discussed in detail. In this section, the researchers discussed similarities and differences in pauses and hesitation in the genderlect of Pakistanis in talk shows.

Table 2. Pauses in Pakistani Male and Female Speakers’ Continuous Speech of 20 to 40 Minutes of the Sampled Talk Shows on PTV World (Shah, 2021)

Filled Pauses (F.P)	Most Common Pauses in Male & Female Speakers’ Speech	Pauses of Male Speakers		Pauses of Female Speakers		Number of Differences		Total Pauses and Hesitations
		No.	%	No.	%	No.	%	
Interjections and Corrections								
Referent Selection	that is = 31 that we consider = 3 as we say = 5 as people think = 1	11	27.5 %	29	72.5 %	F = 18	F = 45%	40
Memory Success	uh = 618 ah = 8 mm-hmm = 1 well = 583	713	58.9 %	497	41.1 %	M = 216	M = 17.8%	1210
Word Approximation	that is = 73 such as = 3 like = 475 in particular = 2	139	25.2 %	414	74.8 %	F = 275	F = 49.6%	553
Exemplification	suppose = 10 such as = 5 let’s say = 2 for example = 34 imagine = 61	63	57%	47	43%	M = 16	M = 14%	110
Repeats	I I = 22 very very = 6 and and = 18 then then = 1 the the = 44 you = 32 that = 36	48	30.2 %	111	69.8 %	F = 63	F = 39.6%	159
TOTAL PAUSES		974	47%	1098	53%	F = 124	F = 6%	2072

Analysis

From Table 2, filled pauses of a number of 11 in Referent Selection, which is 27.5% of the total filled pauses, were used by male speakers. While in the same category the number of 29 pauses which is 72.5% of the total, were used by female speakers. Male speakers used more pauses than female for Referent Selection, the difference of which comes in number 18 pauses, which is 45% of the total. Some of the examples of these filled

pauses used for Referent Selection by male and female speakers in their speech taken from the Table 2 are: that is, that we consider, as we say, as people think. These Filled Pauses (FPs) for Referent Selection (RS) can also be called as Self Addressed Questions (SAQ).

In Memory Success (MS), the number of FPs is 713 which is 58.9% of the total and was used by male and the number 497 pauses which are 41.1%, were used by female. The difference that is more than female, is of male in all types of pauses is this

Corpus Catalogue of Pauses and Hesitations: A Computational Analysis of Genderlect in Pakistani Talk Shows

one as it is found that a number of 216 pauses (i.e., 17.8% of the total filled pauses) was used more by male than female. The example of such pauses as given in Table 2 is: ah, mmm, um, hmm, well, and uh.

In search of FPs for Word Approximation (WA) in spontaneous speech of male speakers, an average of 139 pauses that is 25.2% was found. While in the speech of female speakers, a number 414 pauses that is 74.8% for WA were found. A number of 275 pauses that is 49.6% was the difference found of male from female speakers which means that females use more pauses of word approximation than males. Pauses of this type used by selected subjects as a sample are: that is, such as, like, such that, for example, for instance.

Exemplification is also one of the subtypes of FPs which was considered to be observed and as a result, the FPs of number 63 pauses that is 57% was used by male speakers and the number of 47 pauses was used by female speakers that is 43%. More pauses were used by male speaker than female with a difference of 16 pauses that is 14%. Though not a significant figure, but it is considered as observed for this study. Fillers which were used

for Exemplification are: for example, suppose, such as.

The next type of FPs which were observed in the collected data is Repeats. The number of 48 pauses as Repeats was used by male speakers which are 30.2%. While, female speakers used 111 pauses that are 69% of the total. Female speakers used more pauses of this type than male; the number of difference found was 63 pauses which were 39.6% of the total. For Repeats, both male and female used pauses are: I I, the the, very very, think, think.

Nevertheless, the total FPs found in speech of male speakers was a number of 974 pauses the percentage of which in totality with female was 47%. While, the total FPs found in female speakers' speech was 53%, the number of pauses was 1098. In totality, a minor difference was found in using filled pauses by both male and female speakers, that is 124 filled pauses were used more by female speakers than male speakers, the calculated percentage of which was 6% (Shah, 2021).

Figure 1. Pauses Observed in Pakistani Male and Female Speakers' Continuous Speech

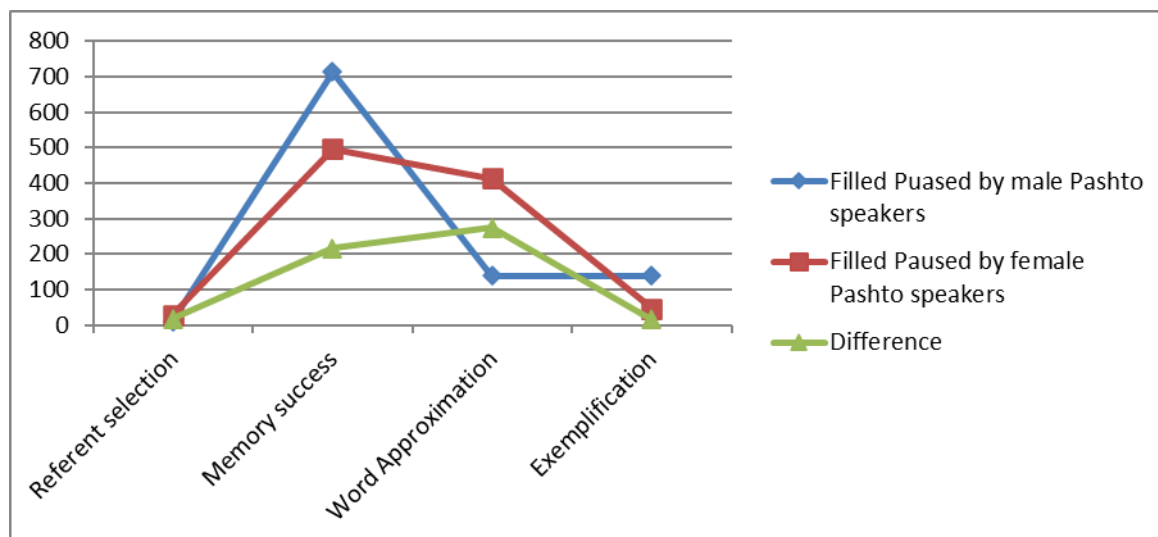


Table 3. Pauses Observed in Pakistani Male and Female Speakers’ Continuous Speech

Filled Pauses (F.P)	Observed Pauses in Male & Female Speakers’ Speech
Interjections and Corrections	
Referent Selection	that is, that we consider, as we say, as people think, you know what?
Memory Success	ahh, Mmm, um, hmm, well, ah, uh
Word Approximation	that is, such as, like, such that, for example, for instance
Exemplification	for example, suppose, such as
Repeats	I I, the the, very very, think, think

3.1.1. Filled Pauses in the Corpus:

This section provides a summary of the corpus in terms of several types of hesitation phenomena identified in the texts.

Referent Selection:

For the Referent Selection, a high frequency of pauses was observed. Pauses like: that is, that we consider, as we say, as people think, taken from Table 3 were used by speakers. In order to buy some time for the selection of an appropriate thing that is related to the context, speakers use these pauses as observed in the transcribed speech recordings. The following extracts exemplify these pauses found for the RS.

In the first example, the speaker prepared himself to purposively select the idea for the addressee to think upon so he used pause of RS in the form of Self Addressed Question (SAQ) “You know what?” for giving importance to his talk.

Likewise, in second example, for choosing an adjective or complementing the subject, a pause in the same manner for RS was used in the form of SAQ (i.e., what?). In the third example, the choice of action to be taken towards the behavior of someone was made by using SAQ for RS.

Memory Success:

Using these pauses of Memory Success, the speaker seems to be able to absorb the subsequent word or phrase for a small period of time. As a result, the grammatical and lexical options are

constrained by the speaker’s usage of it. These pauses were used in speech as given below:

Shahzad Khan: *Well* this morning it’s on your television screens is gonna pop up yes hey come on.

Shazia Hashmi: Twitter you can find us as well this As well this morning without eg.

Shazia Hashmi: Dailymotion on You Tube.

Shahzad Khan: *Well* this morning *well* this morning.

Shazia Hashmi: And the repeat here you can catch us.

Shahzad Khan: By past 11:00 p.m. tonight till the next time one two three good morning.

Salaman Javed: It’s Sophie song with Fatch Ali Khan.

Shazia Hashmi: So let’s do it.

Absolutely, look, we have got to take a break there is more with Alison when we get back. You are now the host of design star next gen. Which is the perfect job for you? And why that is the perfect job?

In the 3rd line, the speaker tries to buy some time in order to speak an appropriate word/sentence in that specific context so he used WELL filled pause of memory success. In the same way, the 4th line also shows the process as mentioned in the explanation of the first example.

Corpus Catalogue of Pauses and Hesitations: A Computational Analysis of Genderlect in Pakistani Talk Shows

Word Approximation:

Word approximation indicates that the speaker thinks and chooses the appropriate word for a description of a thing. The speaker selects a particular word which fulfilled the aim of their description of a sentence. From the transcribed data we can exemplify word approximation in this way.

Shehzad: So has school always been easy for you.

Salman Javed: Yeah I mean I always take a knowledge instantly so I read don't need a whole lot of review of exam prep. For example I can get it instantly.

Shazia Hashmi: Right so do have instant like do you have photographic memory Mike yeah so read too much remember anything I basically yeah and

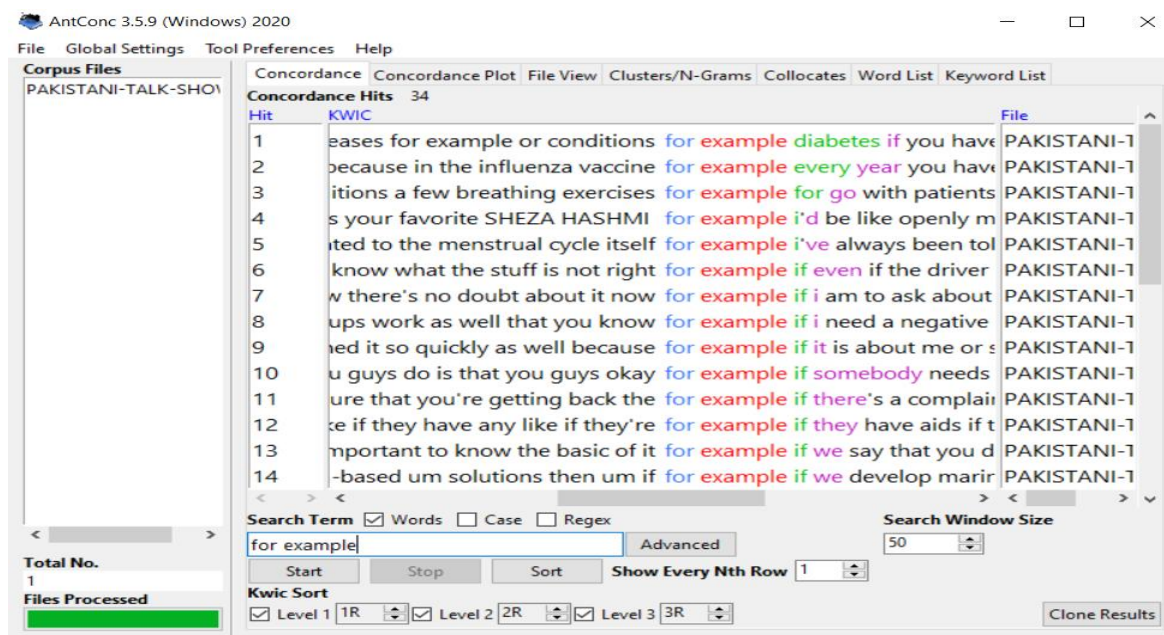
In the first example, the speaker selected/chose the appropriate word from the given table in order to describe the word more accurately.

In the second example, the speaker wants to say the exact and accurate meaning of the sentence. So, the speaker chose appropriate word for the description that it exactly means here.

Exemplification:

The word exemplification indicates that the speaker gave example of what they were saying so the speaker chose the appropriate word from the given table for clarification. The following extracts exemplify these pauses found for the Exemplification.

Figure 2



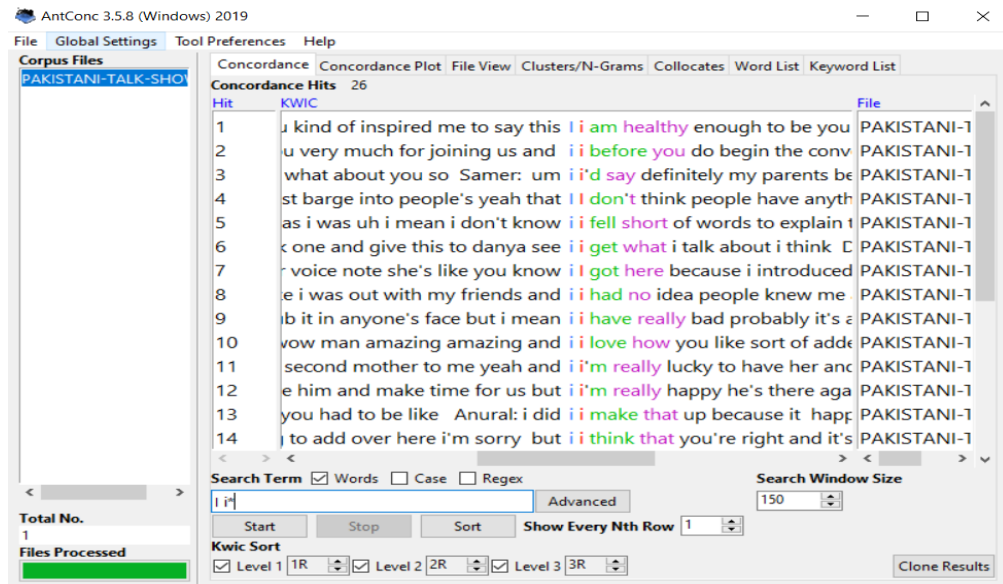
Repeats:

In the example, the speaker tries to understand the listeners by exemplifying their phrase/sentence. They gave solid examples in order to bring coherence and clarify the sentence.

Repeats can vary from a single phenomenon to an extended or stretch of words and sentences that could theoretically be of any length but actually do not exceed four or five words in this corpus. From the transcribed data we can exemplify word approximation in this way.

- 1 Dr. Rashid Aftab: Since 70 percent of the earth is occupied by the uh by the oceans and the sea uh then then in the in earth summit adopted a resolution that each year this day will be observed as a world ocean day number one then why the oceans.
- 2 Sameer: I was able to grow tulips irises and whatnot in Islamabad I was like Islam about the weather is highly possible yeah I was like okay uh is this a fluke or what but then I, I was like no Islam the weather I was really nice people started asking me for my flowers and they asked I, I would be willing to sell them because they were very different from what you were getting in the markets of Islamabad I was like why no.
- 3 Dr. Salim: They will have to show and continue to show aggressiveness in order to keep the control on the valley because otherwise if they start listening to what the the saner voices are saying not only around the world but within India also you know some people who have been quite open.
- 4 Shehzad: That's wonderful I think it's a brilliant idea but there's one more question please and and I think that it's a very pertinent question because this is something which the mothers don't like you know you go uh and you you're actually supposed to buy vegetables.
- 5 Hira: Unless the steps of fifth August are revoked they're reversed which is very difficult I mean you you can't expect a person like Prime Minister Mori to sort of take back all that he has done it.
- 6 Shehzad: Exactly and the part of that that as I mentioned the machinery that we're being bringing to the local market.

Figure 3



- Learning and Verbal Behavior*, 4(6), 469–475.
- Khan, A. Y., Razi, A., Mirza, R., Mazhar, S., Amjad, A., & Shafique, U. (2013). Impact of mass media in Pakistan on social, ethical, and economic grounds. *Int J Eco Res*, 4(3), 1-20.
- Kitzinger, C., & Wilkinson, S. (2017). Referring to persons: Linguistic gender and gender in action – (when) are husbands men? In Geoffrey, R., Gene, H. L., & John, H. (Eds.), 'Enabling human conduct: Studies of talk-in-interaction in honor of Emanuel A. Schegloff' (pp. 189-205). *Pragmatics & Beyond New Series*, 273. Amsterdam/Philadelphia, USA: John Benjamins Publishing Company.
- Köhler, A. (2008) Gender and language. Retrieved from [https://books.google.com.hk/books?id=Wv4YyPTDQWMC&pg=PP1&dq=%C3%B6hler,+A.++\(2008\)++Gender+and+Language.&hl=zh-CN&pg=PP1#v=onepage&q&f=false](https://books.google.com.hk/books?id=Wv4YyPTDQWMC&pg=PP1&dq=%C3%B6hler,+A.++(2008)++Gender+and+Language.&hl=zh-CN&pg=PP1#v=onepage&q&f=false)
- Lakoff, R. T., & Tannen, D. (1979). Communication strategies in conversation: The case of scenes from a marriage. *Annual Meeting of the Berkeley Linguistics Society*, 581-592.
- Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, 14, 41-104.
- Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word*, 15(1), 19-44.
- Mittell, J. (2004). *Genre and television: From cop shows to cartoons in American culture*. New York and London: Routledge.
- Rochester, S. R. (1973). The significance of pauses in spontaneous speech. *Journal of Psycholinguistic Research*, 2(1), 51-81.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Linguistic Society of America*, 50(4), 696–735.
- Shah, T. (2021). A corpus-based comparative study of genderlect in Pakistani and American talk shows [Unpublished MPhil thesis]. Department of English, Abdul Wali Khan University Mardan.
- Sullivan, J. L. (2013). *Media audiences : Effects, users, institutions, and power*. Thousand Oaks, California: Sage Publications.
- Tannen, D. (1990). *You just don't understand: Women and men in conversation*. New York: Ballantine Books.
- Viola, I. C., & Madureira, S. (2008). The roles of pause in speech expression. *Speech Prosody*, 721-724. Retrieved from: https://www.isca-speech.org/archive_v0/sp2008/papers/sp08_721.pdf
- Zellner, B. (1994). Pauses and the temporal structure of speech. In E. Keller (Ed.), 'Fundamentals of speech synthesis and speech recognition' (pp. 41-62). Chichester: John Wiley.