

Spontaneous puns:

The relationship between punning and  
language makeup

27-6-2018

Esther Vullings s4374320

# Inhoud

Inhoud.....	1
Abstract .....	3
Acknowledgements .....	3
1. Introduction .....	4
1.1 Humor and puns .....	4
1.2 Normal word perception and production processes .....	6
1.3 Pun types and language processes.....	10
1.4 Otake and Cutler.....	14
1.5 Phonemic repertoire .....	18
1.6 Stress Shifts .....	19
1.7 Research questions .....	22
1.8 Outline .....	22
2. Methods and data collection.....	23
2.1 Effe Ekdom and Ekdom in de ochtend.....	23
2.2 Types of puns .....	23
2.2.1 Homophone puns .....	24
2.2.2 Embedded puns .....	24
2.2.3 Mutation puns .....	24
2.3 The corpus .....	25
2.4 Procedure.....	27
2.5 Data analysis .....	27
3. General Results .....	29
3.1 Descriptive analysis of the corpus of Dutch puns .....	29
3.2 Source word characteristics.....	30
3.3 Pun type and the MaxSP .....	31

3.3.1 Homophones .....	33
3.3.3 Mutations .....	34
3.3.2 Embeddings.....	34
3.4 Source word preservation and stress .....	36
3.5 Position in word .....	37
4. Differences in source word production .....	38
4.1 Source word characteristics.....	38
4.2 Pun type and the MaxSP .....	39
4.2.1 Homophones .....	40
4.2.2 Mutations .....	40
4.2.3 Embeddings.....	41
4.3 Source word preservation and stress .....	42
4.4 Position in word .....	43
4.5 Difference in data sets .....	43
5. Discussion and conclusion .....	44
5.1 The MaxSP.....	44
5.2 Pun types .....	45
5.3 Phonetic and prosodic makeup.....	47
5.4 Puns and word recognition and production processes.....	49
5.5 Limitations and recommendations for further research .....	50
5.6 Conclusions .....	51
6. Bibliography.....	53

## **Abstract**

The purpose of this study is to investigate the relationship between the act of making spontaneous puns and the way a language is made up, both phonetically and prosodically. The study is partly based on a similar study on Japanese that was conducted by Otake and Cutler in 2013. An analysis was done on a corpus of spontaneous Dutch puns produced by Gerard Eklom, a radio show host. The results of the study suggest that preservation of the source word is the most important factor in the act of punning in both Dutch and Japanese. Phonetic makeup of a language does not seem to affect the type of puns that appear in a language. The disfavoring of stress shifts between source words and target words in Dutch suggests that prosody in a language does affect the act of punning.

## **Acknowledgements**

I would like to thank my thesis advisor Mirjam Broersma for her support and much appreciated advice whenever I needed it. I would also like to thank Sybrine Bultena for being the second reader of this thesis. I would also like to thank my friends and family for their support. Finally, a special thank you goes out to a band called The Maine for keeping me sane with their music throughout the process of writing this thesis. Thank you.

Esther Vullings

# **1. Introduction**

Puns are made and heard by many people and appear in a lot of places like television, radio, print media and even advertising. Puns are a popular tool in advertising and serve an important role in entertaining the reader or viewer (Van Mulken, Van Enschot-van Dijk & Hoeken, 2005). The psycholinguistic processes that underlie this humorous effort are quite interesting. According to several studies, that will be discussed in more detail later, normal word recognition and production processes are the foundation of the action of punning, but phonological structures of languages may also play a role in punning. This paper will focus on the underlying processes of punning and will address the result of the phonological makeup of the Dutch language on puns. Comparisons of results drawn from the Dutch language with a study on Japanese puns by Otake and Cutler (2013) will be made. The Japanese study will be partly replicated with small methodological additions and changes to fit the Dutch language.

## **1.1 Humor and puns**

The processes that underlie the action of punning can be partly drawn from theories of humor in general. Humor, according to McGraw and Warren (2010), requires two contradictory ideas in the same situation at the same time. Anything that is threatening to how someone expects the world to be will be seen as humorous as long as this is also benign. This is what McGraw and Warren named the Benign Violation Theory.

The Benign Violation Theory of McGraw and Warren (2010) suggests that three conditions are necessary for eliciting humor. The situation has to be a violation, this violation has to be benign and these two conditions have to occur at the same time. Puns are an example of this since they usually have two different, and sometimes contradicting, meanings that are brought together at the same time to elicit a humorous effect.

Punning and wordplay, according to Partington (2009), receive, out of all types of humor, the most attention in the study of linguistics. This is quite obvious, states Partington, because puns draw explicitly on wording and language. But how can the term ‘pun’ be defined? What constitutes a pun? And when is a joke considered a pun?

Partington (2009) has the answers. Most theoretical frameworks on puns, he states, agree on the fact that puns involve two different senses or meanings. He adds that puns are not only plays on meaning but also on sound and the resemblance of sound between two words or phrases. Puns do not always play on words only, most of the time they involve entire

phrases or sentences. According to Dynel (2010), not all meanings of the word that's punned with have to be relevant to the joke.

A lot of different classifications of different types of puns have been made over time. Partington (2009), for example, addresses the distinction between 'near puns' and 'exact puns'. This distinction was also made by Miller, Hempelmann, and Gurevych (2017) who state that the latter is commonly known as 'perfect puns' rather than 'exact puns'. An exact or perfect pun draws on sound sequences that are identical to each other, while near puns only resemble each other phonetically.

Another categorization of puns and wordplay is made by Ritchie (2005). He makes a distinction between puns that are self-contained or contextually integrated. Self-contained puns are jokes that do not directly link to the context in which the pun occurs semantically. The context is not necessary to the joke structure for this type of pun. According to Ritchie, these kinds of puns draw more on general knowledge. An example of a self-contained pun is as follows:

(1) Self-contained pun

Ritchie, 2005

What do you get when you cross a murderer with a breakfast food? A cereal killer.

A contextually integrated pun is, according to Ritchie (2005), a pun that draws on context and further linguistic properties to make it a pun. An example of this type of pun is also taken from Ritchie's paper.

(2) Contextually integrated pun

Ritchie, 2005

A shopper is walking along, and a leek falls from his shopping bag to the ground, unnoticed. Another shopper calls out, "Hey! Your bag's leaking!"

The classification of puns that will be maintained in this study will be the one made by Otake and Cutler (2013). They distinguish between three types of puns. The first type is homophone puns, which are words that sound the same and may also be spelled the same but have multiple different meanings. The second type is embedded puns, in this type of pun a word is either inserted into a longer word or extracted from a longer word. The last type is what Otake and Cutler named mutation puns. The source word this type of pun is based on is changed by

inserting, deleting or changing one or more sounds. The decision to keep this categorization has been made to replicate Otake and Cutler's study as closely as possible.

## **1.2 Normal word perception and production processes**

The act of punning, according to Otake and Cutler (2013), draws on normal word perception and production processes. One of these processes, the word production process, deals with several separate processes: the selection of the semantic and syntactic makeup of the word, the retrieval of the phonetic properties of the selected word, the syllabification of the word and actually pronouncing the word correctly (Levelt, 1999).

When a word is selected to be produced by the speaker, semantically and phonetically similar words are also activated in the speaker's mind (Levelt, 1999). When a speaker wants to produce the word 'cat', the error 'rat' is likely to occur, because of the activation of related words. The word 'rat' is semantically related to 'cat' because cat and rat are both animals. But the two words also sound alike. This makes the mistake 'rat' instead of 'cat' more likely than 'dog', a word that is only semantically related to 'cat'. This activation process is much like the process of recognizing words by a listener. These recognition processes will be explained in more detail later.

According to Levelt (2000), a core process in speech production is preparing words from a semantic base. The first step in this process is lexical selection. A person wants to look for a word to say to serve a specific communicative goal. The speaker focuses on a concept and then selects the lemma that corresponds with that concept from their mental lexicon. While selecting the correct lemma, other (related) lemmas are activated as well. After the selection of the lemma, the syntactic properties become available and form encoding can start. When the item is selected, the phonological code of that lemma needs to be accessed by the speaker. In the case of synonyms, both phonological codes get activated in the speaker's mind. The access of the phonological code of a word happens more quickly for highly frequent words than for words that occur in language use less frequently.

After receiving phonological codes, syllabification happens, after which phonetic encoding makes sure the word can be articulated properly and the speaker can produce the selected lemma. This process happens very quickly, but despite this speed, we still only produce an error once every 1000 words (Levelt, 1999).

Indefrey and Levelt (2004) discuss the order in which the different parts of the word production process happen. According to them, the normal word production process starts with activating a lexical concept and selecting that concept for expression. The next stage is

the access to the syntax of the selected word. This is what Indefrey and Levelt also call grammatical encoding. After grammatical encoding of lemma nodes, the phonological code needs to be accessed.

Indefrey and Levelt (2004) also suggest an integration of a person's perception and production processes. This integration has been suggested by the fact that semantically related words can slow down the naming process of concepts. Distractor words that are phonologically related can actually speed up this process. The process can also be affected and disturbed by non-words as distractors. These effects of related or non-related words prove the fact that word perception and production processes are integrated with each other.

After a punster has produced their pun, it needs to be received, segmented, recognized and understood by the listener. This is where speech perception processes come into play. These perception processes also play a role in picking a source word to base a pun on by the punster.

The speech that a listener receives and needs to be recognized is a continuous signal that first needs to be segmented into words by the listener to fully understand what has been said. According to Cole, Jakimik, and Cooper (1980), that is one of the most challenging problems in the study of language perception.

Cole, Jakimik, and Cooper (1980) state that fluent speech can often be parsed in multiple ways and not only one. According to them, puns are often based on this fact that fluent speech can sometimes be parsed in multiple different ways. This is also what makes speech segmentation so challenging. Phonological processes like palatalization, in the example that Cole et al. give in the paper: 'six sheep' and 'sick sheep', can also make sentences ambiguous. In four experiments it was tested how ambiguous and non-ambiguous sentences were perceived by listeners. Participants were told to react when they heard a mispronunciation in the story they were listening to.

The four experiments showed that listeners can use context information to build up expectations of what a speaker is going to say. This expectation is the basis for a definitive parsing of the sentence. The expectation of what will be said can also be based on words that occur just before the ambiguous part of the sentence. Cole, Jakimik, and Cooper (1980) conclude that thematic, semantic and syntactic information can be used to segment a continuous stream of language into words.

Not only does speech need to be segmented, words need to be recognized and processed as well. A lot of studies have given more insight into these processes. One of those studies was done by McQueen, Cutler, and Norris (1994). It has been stated that words will be



recognized at the point they become unique and can no longer be another word. Short words, McQueen, Cutler, and Norris state, only become unique after their offset. They say that it is unlikely that word recognition happens in such a sequence. Three experiments proved that, instead of sequential recognition, competition between several word candidates is a far more likely process in word recognition.

Several models of word recognition have been developed based on what we know of the process. There are two different ways to look at the process of word recognition: the Cohort Model and the Neighborhood Activation Model. This is explained by McQueen and Cutler (1992). Both models integrate phoneme repertoire and vocabulary structure in a different way. One thing they have in common is the fact that embedding in words plays an important role. The Cohort Model, according to McQueen and Cutler, recognition is based on how many other words share the initial portions of the word. The Neighborhood Activation Model, however, depends on how many words resemble it at any point in the word.

According to Marslen-Wilson (1987), normal spoken word recognition processes, consist of three separate functions. These functions are access, selection, and integration. Access concerns the mapping of the received speech information to a representation that is available in the mental lexicon of the listener. Integration provides a basis to a higher level representation. The final function selects the word form in the mental lexicon that represents the speech input best.

Even before the full phonetic information is available in utterance contexts. Marslen-Wilson (1987) states, a word is already recognized by a listener. For a word in isolation to be recognized, the listener needs a little more time to process it. Content words are recognized very early while shorter function words and less frequent content words are not recognized as early. This early selection of words is of importance in the theory of multiple access. Multiple access happens when multiple potential candidates of a word are activated in a listeners mind when hearing phonetic input. For puns and wordplay, the activation of multiple potential candidates is an important factor for understanding a pun and for the punster to select a word to pun with. When understanding either embedded or mutation puns, more than one word needs to be activated and for homophones, the multiple meanings of the word need to be available to the listener.

The activation of multiple candidates is similar to lexical competition where identical words have different meanings: homophones. Dahan, Magnuson, Tanenhouse, and Hogan (2001) conducted an eye-tracking experiment to find evidence for this type of lexical competition. Participants fixated on the target word slower when the onset of a target word

came from a word competing with the target word than when the onset came from a non-word. Competing meanings can slow down activation of target words in the brain of a listener.

It has also been reported that lexical competition could facilitate a faster recognition of a word. It has however been proven by Rodd, Gaskell, and Marslen-Wilson (2002), like the earlier discussed study by Dahan et al. (2001), that ambiguity in meaning can actually delay the recognition of either meaning of the ambiguous word. Multiple related senses of a word, like 'twist', can, however, facilitate faster recognition. This could have a detrimental effect on the recognition of puns.

Dynel (2010) studied the relationship between language processes and the act of punning. According to Dynel, a punster will (when the moment allows it) reuse witticisms stored in their mental lexicon to amuse the listener they are conversing with. Said listener has the tendency to not expend a lot of cognitive processes. A hearer will use as little cognitive load as possible. This seems to be opposite to what understanding a pun contains: a little extra mental effort. A hearer thus has to use more mental effort to understand the ambiguous meanings that a lot of puns usually contain.

This ties in with the Relevance Theory of Sperber and Wilson (2004). Under normal circumstances, a listener will be satisfied with understanding one meaning that logically fits the context in which something has been said. For puns, this is different. Both meanings of a homophonic word are important and need to be recognized by the listener.

Research on the word recognition process while using puns as data was done by Otake and Cutler (2001). They state that when a person hears a speech input, multiple candidate words are activated that are partially or fully compatible with said input. It is common for puns to involve distortion of words. Most puns draw on homophony or mutate one single phoneme of a word. Adding and deleting sounds do also occur. When someone hears a pun like that, for it to be successful, the actual word as well as the intended base or meaning need to be activated. If not, the joke will not be received well by the listener.

By means of an analysis of a corpus of Japanese puns, Otake and Cutler (2001) concluded that punning involves the basic and normal processes of word recognition. The analysis also showed that most puns resemble the target entirely, except for a single phoneme. This is remarkable because speakers of Japanese are more aware of moras than phonemes. Even with moraic awareness, words are most often distorted by only one single phoneme. This should thus be a fact across multiple languages.

### **1.3 Pun types and language processes**

It has already been established that the act of punning draws on normal language perception and production processes, but different types of puns require different types of processes and can have different effects on word production and perception processes. Otake and Cutler (2013) did a study on puns in Japanese and distinguished three different types of puns: homophones, embeddings and mutations. Homophones are two words that sound identical but carry different meanings. In embedded puns, a target word is either embedded in or extracted from a source word. Mutation puns change one or more sounds in the source word to create the target word (the pun).

Puns are mostly based on similarity of words that sound either partly or fully the same. Hahn and Bailey (2005) studied what it is that makes words sound similar to us. They write that it is widely stated that the similarity between two similar words is based mostly on the same phonemes that are located in the same position of a word. This is what makes two words sound similar to us. A series of five experiments have confirmed that these models and statements are true.

The first mentioned type of pun that Otake and Cutler (2013) studied is the category of homophone puns. Rodd, Gaskell, and Marslen-Wilson (2002) have studied recognition of these semantically ambiguous words with multiple meanings (homophones). Several studies that were done before the one by Rodd et al. have stated that ambiguity between multiple meanings of words can speed up visual lexical decisions. Rodd, Gaskell, and Marslen-Wilson conducted three experiments testing this statement. In the experiments, the participants had to do a lexical decision task for ambiguous and non-ambiguous words. All three experiments showed an ambiguity disadvantage rather than an advantage. Words with multiple senses did, however, speed up lexical decision. This could, according to Rodd, Gaskell, and Marslen-Wilson, explain why other studies found an advantage of ambiguity.

The disadvantage for semantically ambiguous words on lexical decision is, according to Rodd, Gaskell, and Marslen-Wilson, also in line with speech perception models that contain competition between different words that can be activated. The interference of different meanings of the same word delays the recognition of the word in question. These findings are evidence for a competition based speech recognition and word activation in the word recognition process.

Homophonic words do, however, not share the same frequency in language use. Research about this frequency of appearance has been done by Caramazza, Costa, Miozzo,

and Bi (2001), Gahl (2008) and Sandra, Frisson, and Daems (1999).

The study by Gahl (2008) confirmed that words that appear more frequently in use tend to shorten in pronunciation. This is, according to Gahl, because how more often we do something, the faster we can do it. She concludes that this is the reason ‘time’ and ‘thyme’ are not homophones. They share a phonological encoding but not a lemma in the language speakers mind.

Durational differences between homophonic words were also found by Warner, Jongman, Sereno, and Kemps (2003). They state that speakers can produce small differences in the duration of words that are believed to be homophones. These durational differences are named ‘sub-phonemic durational differences’ by Warner et al. An experiment revealed the fact that those differences can be used by listeners to distinguish words with different underlying forms. Two words that are homophones can be pronounced differently on an almost inaudible sub-phonemic level. When making puns, this effect can lead to a quick understanding of the two ambiguous meanings a punster is trying to convey.

Caramazza, Costa, Miozzo, and Bi (2001) found that the frequency of a specific word can cause latencies on word naming tasks. Not only does this effect happen for spoken language, homophone frequency can also affect written language. Sandra, Frisson, and Daems (1999) found that spelling errors of homophonous verb forms in Dutch are determined by the frequency of the word in question. The more frequent a word is, the fewer errors are made. The results of all three studies support the hypothesis that the different meanings of homophones do not share the same word representation but both have independent representations in a language users mental lexicon.

The second category of puns that Otake and Cutler (2013) mention is the category of embedded puns. As already explained, this is a type of pun where one word is either inserted in or extracted from another word. A hearer has to spot a word in another word to understand the pun at hand. Research on embedded words by Norris, McQueen, and Cutler (1997), resulted in a model of language processing that was named the Possible Word Constraint. This is a theory of speech segmentation that is compatible with completion based language perception. The recognition of a stream of continuous speech is, according to the study, subject to constraints of what may possibly be a word in the language that is spoken. Two word spotting experiments were carried out by Norris et al. Participants heard non-words and had to spot words that were embedded in the prime word. In the first experiment, the participants heard bisyllabic and monosyllabic non-words that had another word embedded in it. Examples of the words were ‘vuffapple’ and ‘fapple’. One condition was a non-word

containing a possible other word like ‘vuffapple’ and the other condition contained only one phoneme that could not be a word like ‘fapple’. Reactions to the word spotting were more accurate and faster in the condition that contained a possible word than in the condition that contained an impossible word. The second experiment cross-spliced the data across conditions and again confirmed the hypothesis tested in the first experiment. Words are easier to spot in possible contexts than in contexts containing words that are not possible in the English language.

More research on word frequency was done by McQueen and Cutler (1992), who analyzed the frequency of embedded words and the position in which two words resemble each other. It is concluded that most polysyllabic words begin with other words. Most embedded words in English have a word-initial embedding (94% of the analyzed words). It is not unreasonable to hypothesize this will also be the fact in a Dutch corpus of spontaneous puns. McQueen and Cutler conclude that the recognition of these longer embedded words will involve the rejection of other shorter words.

According to Bowers, Davis, and Hanley (2005), shorter words are often embedded in larger, carrier words in speech. In a series of experiments, it has been examined whether those shorter words are activated when a listener hears the bigger word. The experiments resulted in evidence for the hypothesis that smaller parts of a larger word are in fact activated when someone is presented with the larger word. This happens in initial, medial and final positions of a word and the words are activated at both the level of form and meaning.

Isel and Bacri (1999) have also studied the activation of smaller words embedded in either initial or final position of larger words. The goal of the study was to know whether lexical access is triggered only at the onset of a word or at any position in the auditory input. According to Isel and Bacri, the existing Cohort model of speech recognition does not predict final embedded words to be activated, contrary to what Bowers, Davis, and Hanley (2005) have concluded. Initial embedded words will, according to Isel and Bacri, partly be activated only at the onset of the word.

To try and confirm this hypothesis, Isel and Bacri (1999) studied the access of initial-embedded words in their first experiment. Forty native speakers of Parisian French had to decide whether the stimulus they received was a word or a non-word. The second experiment tested the activation of the final-embedded words. It turns out that final-embedded words are sufficiently processed by listeners to show priming effects. Isel and Bacri conclude that lexical access is triggered in at least two points, the beginning of a word and at a median point. The fact that final embedded words are also processed makes final embedded puns

possible for punsters and accessible to listeners. Final embedded puns can and should thus appear in the act of punning in both extracted and inserted form.

Another study on embedded words was done by Vroomen and De Gelder (1997). The results of this study support the theory that metric segmentation and lexical competition operate together. In a first experiment, participants heard a Dutch word like ‘framboos’ (raspberry) in which the word ‘boos’ (angry) is embedded in the word end. In a lexical decision task, the participant saw words that were associated with the word embedded in the word endings.

Vroomen and De Gelder (1997) conclude from the results of this experiment that the words embedded at the end of the source word are also activated as possible candidates, which is compatible with other studies. A second experiment tried to replicate these results, but with words that have a weak-strong pattern like ‘vervoer’ (transport). The first syllable of the input words was an unstressed prefix followed by a stressed syllable that is also a word. This experiment showed the same effect as the first experiment: embedded words are activated.

The third experiment that was conducted by Vroomen and De Gelder (1997) used words with only a single syllable but still had another word embedded. Does ‘zwijn’ (swine) also activate the word ‘wijn’ (wine)? This was also tested with beginning embedded words. No effects were found in this experiment. Only words embedded in multi-syllabic words are activated when heard. This would mean that embedded puns are most likely to consist of more than one syllable. This means that inserting ‘boos’ in ‘framboos’ in a Dutch pun would be preferred over embedding ‘wijn’ in ‘zwijn’.

The third and last pun type that Otake and Cutler (2013) researched and that will be included in this study is the category of mutation puns. A mutation pun is a pun that mutates the source word, preferably by as little as only one or two phonemes. The way words can be mutated has been researched by Van Ooijen (1996), but instead of word spotting experiments (that were carried out by Norris, McQueen, and Cutler, 1994), participants had to reconstruct words that had been changed by mutating sounds.

In Van Ooijens (1996) experiment on the mutability of words, participants were given non-words like ‘eltimate’ and had to reconstruct an existing word like ‘ultimate’ by changing the vowel, or ‘estimate’ by changing a consonant. Participants had to press a response button as soon as they could identify a real word that resembled the stimulus. They were then asked to present this word verbally. If there was more than one word, they were asked which one came to mind first. The results show that vowels were easier to substitute in a word than

consonants. Even when participants were explicitly told to pay attention to consonants, responses based on vowels occurred. Vowel mutations outnumbered consonant mutations. Most of the vowels were strong and unreduced vowels. These results suggest a mechanism that deals with expected vowel variability in continuous speech signals.

Good puns, as stated by Otake and Cutler (2013), preserve as much of the source word that the pun is based on as possible. Mutations of one or two sounds are thus preferred over mutations of more sounds. This was also found in a study on Japanese puns by Shinohara and Kawahara (2010). Speakers tend to minimize differences between source words and target words, even when intruding syllables. These results show, according to Shinohara and Kawahara, parallels between puns and natural language processes.

With the results of the studies I have mentioned in mind I hypothesize that a punster making mutation puns would rather make changes in a vowel than in a consonant since this type of mutation is easier to make and to understand. Expected vowel variability would make it easier to come up with and recognize puns with a mutation of the vowel. Changing fewer sounds rather than more sounds will also be preferred for punning.

According to Kawahara and Shinohara (2011), in Japanese imperfect puns, people make sentences by using words or phrases that sound similar. When creating these imperfect puns, speakers avoid mismatches between the two similar sounding words in prominent positions. This happens in initial syllables and long vowels as well. These effects had been found in normal speech and phonology but two experiments by Kawahara and Shinohara have proven this to be true for the formation of puns as well. This is in line with what Otake and Cutler (2013) state about their Principle of Maximal Source Preservation (MaxSP), which will now be explained in more detail.

## **1.4 Otake and Cutler**

The current study on Dutch puns will mostly be based on, and will partly be a replication of research on Japanese puns by Otake and Cutler (2013). This study has been mentioned multiple times already but will now be explained in more detail. Otake and Cutler put together and studied a corpus of spontaneous Japanese puns that were all produced by the same speaker. Puns and wordplay can give us insight into several psycholinguistic processes that have to do with language production and perception. Issues that arise with this type of research, they state, appear in reliability. Firstly, puns that appear in entertainment are usually well thought out in advance. Secondly, the people that collect puns may only record puns that are memorable. This is why the study by Otake & Cutler (and the current study) was done



using a corpus of spontaneous puns that were all made by the same speaker. The data came from a Japanese radio show where the host interviews people and he looks for a source word in the answer that a person gives and bases his puns on that word.

Conclusions on puns that have been drawn from written data suggest that the best puns are the ones that stay closest to their source word. This phenomenon is what Otake and Cutler (2013) named the Principle of Maximal Source Preservation (MaxSP). With this knowledge, the question whether this principle is only a criterion that is used for the quality of puns, or whether this principle also plays a role during the spontaneous production process of puns and wordplay, was explored.

This question was answered by using the before mentioned corpus of puns made by the Japanese radio show host Dokumamushi Sandayuu (D). An important feature of the corpus used is the fact that the puns are all spontaneous and thought of on the spot, during a conversation. None of the puns have been written in advance. This gave the researchers a view of the punning process in action.

The process of punning draws on natural speech production and speech perception processes. Otake & Cutler (2013) made a list of steps involved in the process of making a pun. The first step is recognizing what the interlocutor has said and evaluating a possible source word for punning. The goal is to make an amusing and funny utterance. The audience needs to laugh and be amused because of this, so the speech perception of the audience plays a role as well as the speech perception of the speaker. Once a transformation of the source word has been chosen, it needs to be uttered by the punster.

The process of making a pun is, according to Otake and Cutler (2013) an overlap of speech perception and speech production processes. Speech perception starts with acoustic and phonetic input and generates the meaning associated. Speech production starts with the meaning and generates the phonetic output to convey said meaning. Making a pun involves an overlap of both of these processes.

When listening to speech, all words that are compatible with a signal will be activated. Words that are partially compatible will be temporarily activated by a listener. This activation of word forms will make homophones readily available and this is compatible with the MaxSP principle. The discarding and maintaining of these competing word forms are, to some extent, under the control of the listener. Otake & Cutler analyzed the corpus with these processes and the relationship between pun and source word kept in mind.

The Japanese corpus was collected by recording segments from a Japanese radio show in which the host interviews people and looks for source words to pun with. This method



yielded more than 400 recordings in total. The recorded materials were looked at and puns were saved as separate files and were then transcribed in Japanese. The definition for ‘puns’ that was used to do this was the following: ‘word transformations, with a source word spoken by an interlocutor and a pun word produced by D.’ Puns where the source word was not part of the utterance of the interlocutor were excluded. 294 single-word puns were collected using this definition. The puns were then each coded as one of the following three pun types:

1. Homophone puns: the source word and the pun are identical in form
2. Embedding puns: the source word and the pun have a part-whole relationship. The pun is either extracted from a longer source word, or the source word is inserted into a longer pun.
3. Mutation puns: the source word and the pun are not identical in form, but part of the source is changed to form the pun.

After the coding, the puns, pun properties source words and context were statistically analyzed. The first observation made was the fact that all of D’s puns include nouns and the utterances where puns were chosen from were quite short. According to the MaxSP, it has been predicted that the most preferred pun would retain the integrity of the source word and should thus be a homophone. Languages like Japanese, that have a simple phonological structure, tend to contain a lot of homophones. But it is not the case that all words are homophones. 36% of all Japanese words are homophonic, and this would thus be the maximum amount of homophone puns expected in the corpus. This number holds up for the data in the Japanese D-corpus. The corpus contains 105 homophones, 96 embeddings, and 93 mutations.

The puns were also analyzed for the opportunity they offered for an alternative pun type. The MaxSP predicts that mutations will occur when source-preserving types are not an option. But 81 of the 93 cases would allow another option than the mutation that D chose. Of embeddings with insertion, 47 of the 51 cases could become another word with the change of only one sound. Homophones were not only chosen because no alternative was possible, of 105 cases, 102 words allowed for another word to be formed by changing a single sound and would be able to be a different type of pun. Source-preserving options are chosen significantly more often than non-preserving options, but the non-preserving options outnumber the source preserving options. The choices do show a preference for homophones where that is a possibility. Embeddings are preferred over mutation puns. This preference is in line with what the MaxSP predicts.

A factor in embedding and mutation is the question of what part of a word can be changed and what cannot be changed. Otake and Cutler (2013) have analyzed this. There are, according to them, three levels of source word preservation. The first level (a) is a preservation of the abstract phonological structure. The second level (b) is the segmental structure and the third and last level (c) is both the segmental and prosodic realization. In English puns, the last level of preservation is the most preferred type of preservation. Japanese could, according to Otake and Cutler, differ from English because it is a language has a prosody that is based on pitch accent. Words with accent have only one possible prosodic structure. In the corpus, it was seen that puns where word length is preserved, accent is also preserved. This is not the case for embeddings that do not preserve word length. Accent pattern is not a criterion for preservation, but source words can be preserved by ensuring maximal segmental overlap and not by matching the full prosodic and phonetic realizations.

Otake and Cutler (2013) state that not everything in a word can be changed to create a mutation. Minimal changes to the source word are preferred over more drastic changes. The corpus shows a preference for submoraic changes over moraic changes. This means that in 71 cases (of 93) only a single phoneme is changed to create the pun. This is the case for all word positions (initial, medial and final). Word-initial positions are however most likely to be changed in a mutation pun and initial overlap is most likely to be used to create an embedded pun.

The analysis suggests that the selection of puns by D in this corpus is meaning-driven, which leads to a bias towards nouns as source words. The punning is also driven by what was called the MaxSP by Otake and Cutler (2013). This is in line with the hypothesis. The best pun is a homophone because the source word is completely preserved in a homophonic pun. These homophonic puns are recognized by the language perception system because the system activates all words that are compatible with the input. Subsequent processing will determine which of the different words with the same phonological representation was the word in the input. *Banned* may not be immediately ruled out in the context of ‘I play in a band’. This is also the case for embedded puns. Hearing the word ‘cap’ will also activate the word ‘capture’.

The nature of the perceptive and productive language system is a consequence of the structure of a language’s vocabulary. Languages have many words but only a relatively small inventory of phonemes to construct those words with. The mean phoneme repertoire is about 30, Japanese has less than that (25), English has more (46) and Dutch has a little more

phonemes than the average (about 40). The smaller the number of phonemes a language has, the longer words are likely to be and the more embedded forms a language has. Homophony tends to be most likely in languages with smaller phoneme inventories. The question then arises whether homophony plays a smaller role in Dutch puns than it does in Japanese puns since the phoneme inventory is bigger. A first sub-question in this paper will thus be: Does the Dutch corpus contain fewer homophones than the Japanese corpus when comparing different categories of puns?

In English, stress has a different effect on puns than in Japanese, Otake and Cutler (2013) state. Puns that shift stress are not a favorable option in English while in Japanese, stress does not play such a big role. Changing a vowel in a stressed syllable in English can disrupt the word recognition process. Dutch is a language that is more similar to English than to Japanese. In Dutch, stress differences even create words that seem homophones at first, but are not really, because of stress. It is not unreasonable to think that Dutch puns will also differ from Japanese puns in this respect. The second sub-question is thus: do speakers of Dutch avoid stress shifts when making puns?

The study, by Otake and Cutler (2013), conducted on Japanese data has shown the underlying process for successful punning. The normal process of word production and perception facilitates wordplay. Punsters will apply the MaxSP to make the best puns that work well. A question that arises is whether this conclusion will hold for languages other than Japanese as well. The main question that will be explored in this paper is: Do Dutch speakers apply the MaxSP to make puns (mostly homophones)?

Even though Dutch and Japanese are very different, have very different phoneme repertoires and treat stress differently, it is still reasonable to predict the MaxSP to also apply for puns in Dutch. The more a pun shares with its source word, the better it is received and understood. Even with the phonological differences in mind, I want to hypothesize that the category homophones will be the biggest category in a Dutch corpus as well as in the Japanese corpus. The best pun is a homophone.

### **1.5 Phonemic repertoire**

Not only language processes can affect homonymy, homophony, embedding, mutation and wordplay in general in a language. As I have already touched upon in the previous paragraph, phonetic makeup also plays a role. Phonetics are not critical for wordplay but can affect preferences of word choice in a language. As I have already mentioned, and many studies have established, humor in punning comes from ambiguous meanings of similar (or identical)

sounding words. The possibility of making puns with similar sounding words correlates with the size of a phonetic lexicon in a language, according to Otake and Cutler (2013).

In the study on puns, by Otake and Cutler (2013), it was concluded that the best pun is a homophone. This is also the biggest category in their corpus of puns. Japanese has, however, a smaller phoneme repertoire than Dutch has. A phoneme repertoire can have an effect on homophony in a language but can also affect word recognition. According to Cutler, Norris, and Sebastián-Gallés (2004), a language containing a larger amount of phonemes allows for more and shorter words to be formed and allows for a reduced embedding of shorter words within longer words.

Costa, Cutler, and Sebastián-Gallés (1998) compared the Dutch phoneme repertoire to the smaller Spanish phoneme repertoire and their effects on language processing. Languages differ in the types of demand their phonology makes on the listener. Spanish speakers showed asymmetry in recognition of phonemes while Dutch speakers showed symmetrical effects for vowels and consonants. This is in line with the distribution of consonants and vowels in both languages. The phonological makeup of languages seems to play a role in the processing of those languages.

As I have already mentioned, phonological repertoires do not only affect word processing, they also affect possibilities in the language. Languages with a larger phoneme repertoire allow for more words, a reduced amount of embedding and less homophony. Japanese, according to Otake and Cutler (2013) has a smaller than average (the average is 30) phoneme repertoire while the Dutch language contains more phonemes than average. However, no language resists homophony at all so it is expected that the category of homophones will not be entirely absent in Dutch.

## **1.6 Stress Shifts**

The way a language treats stress can also lead to differences between languages when it comes to punning. Some spontaneous puns are made by changing the stress of a word to either make another word, mutate a word or create an embedding. In the study of Japanese puns, by Otake & Cutler (2013), it was mentioned that in English stress shifting puns are not favored. In the Japanese language, stress does not play a big role. Because Dutch is more similar to English than to Japanese, there are reasons to hypothesize a similar effect in Dutch: disfavoring of stress shifts in puns. This would mean that a corpus of spontaneous puns should contain less stress shifting puns than stress preserving puns.

The similarity between Dutch and English stress patterns and lexical access was studied by Cooper, Cutler, and Wales (2002). According to this study, English and Dutch both have stress rhythm based on strong and weak syllables. Another similarity of both languages is that in both languages stress tends to fall on the first syllable of a word. These statements and the similarity between Dutch and English support my hypothesis that, like in English, stress shifts are not favorable in puns.

An example of stress shifting in Dutch is homophones that are not really homophones. They only seem that way, but a different stress pattern changes the meaning of two words that further share all phonological features. Cutler and Van Donselaar (2001) mention the example 'voornaam' which has stress on the first syllable in the noun meaning 'first name' and has stress on the second syllable in the adjective meaning distinguished/respectable. A sentence like 'Hij was niet vanwege zijn voornaam voornaam.' could be a pun, or could not be a pun at all. This is why Cutler and Van Donselaar researched what is activated for a Dutch listener when confronted with these 'homophones'. The results of the experiments suggest a stronger contribution of segmental information than from supra-segmental information like stress. Mismatches that only consisted of a single segment resulted in similar delayed responses to the control group while supra-segmentally mismatched words did not. A word stressed as MUzee facilitated 'museum' significantly less good than muZEE, which is the correct stress pattern for 'museum'.

Prosodic information can be used to facilitate spoken word recognition, according to Cutler and Clifton (1984). This is concluded after a series of experiments that Cutler and Clifton carried out. Misstressed words are harder to identify than words that are stressed correctly. Prosodic structure of a sentence can also give information about the stress in upcoming words. This leads us to believe that prosodic information is used in the word recognition process. These studies by both Cutler and Clifton (1984) and Cutler and Van Donselaar (2001) are of course not the only evidence for this statement.

Another study that shows evidence for these claims is a study by Cooper, Cutler, and Wales (2002) that also states that stress can be used for spoken word recognition. Stress information that does not match can remove competition of the word that mismatches. The experiments lead to the conclusion that stress cues are used in English for word recognition. Hearing 'admi-' with initial stress activates 'admiral' more than 'admiration'. The disadvantage of misstressed words in recognition can be a reason for preferring puns that do not shift stress.

Another study on stress and spoken word recognition was carried out by Van Leyden

and Van Heuven (1996). An experiment using the gating paradigm was done. In a gating experiment, the participant is repeatedly presented with a word of which the presentation time increases. The participant then has to guess what word is being presented. According to Van Leyden and Van Heuven, this method can simulate the word recognition process. The stimuli that the Dutch and English speaking subjects were presented with in this experiment were correctly stressed and misstressed words. For the misstressed words, the stress shifted to a syllable that does not normally receive stress without changing the quality of the vowel. The experiments showed that misstressing a word impairs the recognition of that word. Front-shifting has a bigger effect on recognition than back-shifting and the effect turns out to be bigger for Dutch than for English. According to Van Leyden and Van Heuven, the results are evidence that rhythmically different word candidates are narrowed down in the word recognition process by using cues from lexical stress.

Research on homographs with different stress patterns by Small, Simon, and Goldberg (1988) showed the importance of stress in English word recognition. Homographs, like homophone puns, are words with the same spelling (or sound) that have multiple meanings and/or origins. Earlier research has already established that miss-stressing a word has an effect on lexical access. Small et al. (1988) looked at differences in word recognition when altering stress for homographs and non-homographs in a phoneme detection task.

Participants heard sentences with a correctly stressed word (John needed to conVERT) or sentences with an incorrectly stressed word (John needed to CONvert). In the non-homograph condition, sentences contained a correctly stressed (PEAnut) or an incorrectly stressed word (peaNUT) that was not a homograph. The reaction times (RT) of the participants were measured for all conditions.

For non-homographs, RT to a target phoneme was faster for correctly stressed words than for incorrectly stressed words. There was no direct access to an incorrectly stressed word in a listener's mental lexicon. For homographs, no clear-cut difference in RT was observed. Even when stressed incorrectly, the word was an existing word that was accessible to the listener.

Whether the stress shift disadvantage can be the reason for a preference for stress preserving puns is not certain since shifting stress can also lead to a completely normal homophone with only a different stress pattern than its source word. However, word frequency can have a negative effect on word recognition. Because of this reason, it is still believed that stress shifts in words would not be preferred when making a pun.

Puns could thus still be successfully understood by a listener, even when stress shifts

occur. The only condition is that the stress shift needs to result in an existing word that is accessible in the listener's mental lexicon. However, because puns with stress shifts are disfavored in English puns, there are reasons to hypothesize that this will also be true for Dutch puns. I want to argue that it can be expected that puns with stress shifts will not occur as much in a Dutch corpus of puns than puns that preserve stress patterns.

### **1.7 Research questions**

In this chapter, I have discussed some literature that has led to several questions and hypotheses. To summarize, the main research question of this study will be:

1. Will the MaxSP hold for the Dutch corpus as well as for the Japanese corpus?

The sub-questions that follow the main question are as follows:

2. Does the Dutch corpus contain fewer homophonic puns than the Japanese corpus when comparing different categories of puns?
3. Do speakers of Dutch avoid stress shifts when making puns?

It has been hypothesized that (1) the MaxSP will be true for the Dutch corpus as well. I also want to argue that it may, however, be true that a system with more phonemes generates fewer homonyms and homophones. I have hypothesized, based on the mentioned research, (2) that Dutch will generate fewer homophonous puns than Japanese since the Dutch phonetic system is larger than that of Japanese. My last hypothesis is that (3) puns that shift the stress pattern of the word will not occur in Dutch as much as puns that retain the original stress pattern of the word.

### **1.8 Outline**

The stated hypotheses in this chapter will be tested by means of analyzing a corpus of spontaneous puns. In section 2, the methodology will be explained in more detail. Section 3 will discuss the data from the corpus and the statistical analysis that followed. Section 4 will discuss more results that were drawn from analysis of the corpus and will compare those to section 3. Section 5 will draw conclusions and discuss some particularities and limitations of this study. Section 6 lists all sources that have been used in this study.



## **2. Methods and data collection**

The research questions that were introduced in the previous section will be answered by analyzing data from a corpus containing a set of puns. The corpus consists of spontaneously produced puns by one speaker of Dutch. The data in the corpus have been collected by listening to a Dutch radio show and annotating each pun that was noted. This procedure will be explained in more detail later in this chapter.

### **2.1 Effe Ekdom and Ekdom in de ochtend**

To replicate the Japanese study by Otake and Cutler (2013) as close as possible, the Dutch corpus was based on spontaneous puns produced by a Dutch radio show host. The Japanese corpus contained spontaneous puns by a Japanese radio show host. The Dutch radio shows that were used for the collection of spontaneous and natural language data were Effe Ekdom and Ekdom in de Ochtend. Effe Ekdom was a radio show on 3FM that aired on weekdays between 12.00 and 14.00. The show was first aired in 2010 and last aired in 2015. From 2015 until 2018 Gerard Ekdom hosted a similar show on NPO radio 2: Ekdom in de Ochtend. The host, Gerard Ekdom, is well known as a punster. He is known to produce a lot of spontaneous puns on radio and television and that is why he was a great candidate to base the corpus on.

Because the radio show includes calling listeners and conversing with them and the co-host, Ekdom receives enough input to base his puns on. The show was chosen to match the Japanese radio show that was used in the study by Otake and Cutler (2013). Only puns that have a source word that was produced by either Ekdom himself or someone else were included in this study. Puns that happen without a clear source word were excluded from the corpus. Puns with a source word in the conversation are spontaneously produced and not thought out, scripted or written before the show. Spontaneous puns give a lot more insight in language perception and production than puns that have been written in advance. For puns that are spontaneously produced it is easier to assume the speaker thought of them himself. The fact that all puns are produced by one speaker is important for the reliability of the research and the validity of the conclusions that will be drawn from the data.

### **2.2 Types of puns**

The typology of puns that was used in this study is based on the typology introduced and used in the study by Otake and Cutler (2013). In their study, they distinguish three types of puns (homophones, embedded puns, and mutations) that will now be explained in more detail. Examples of all pun types in English (from Otake and Cutler, 2013) can be found in table 1.



Examples of the Dutch puns as produced by Ekdom divided by pun type can be found in table 2. Source words and target words are marked with bold text in table 2.

### 2.2.1 Homophone puns

A homophone is a word that sounds like another word (or set of words) with a different meaning. A homophone pun is identical (in either sound only or both sound and form) to its source word. Homophone puns are the type of pun that resembles the source word the closest of all pun types. Homophonic puns have been called ‘perfect’ puns by Miller, Hempelmann, and Gurevych (2017) and ‘exact puns’ by Partington (2009) because of this identical relationship to their target word. Table 2 shows examples of homophonic puns (examples 1A and 1B).

### 2.2.2 Embedded puns

Embedded puns have a part-whole relationship with their source words. A source word can be either inserted into a newly formed word or a new word can be extracted from the source word. Both types of embedded puns can be found in table 1 (examples 2A and 2B). Example 2A is an extracted pun and example 2B is an inserted embedding.

### 2.2.3 Mutation puns

Mutation puns include a mutation or transformation of part of the source word. One or more sounds are deleted, changed or added to create a pun. The source word and target word (pun) sound similar but sound by no means identical. These types of puns could be called ‘imperfect’ or ‘heterophonic’ puns (Miller, Hempelmann and Gurevych, 2017). Examples of mutation puns can be found in table 2 (examples 3A and 3B).

**Table 1.** Examples of each type of pun in English (Otake and Cutler, 2013).

(1)	A	Homophone	I sing with a <b>band</b> . Why were they <b>banned</b> , were they that bad?
	B	Embedding: insertion	I’ve lost my cap and I need a new one. I hope you capture one!
	C	Embedding: extraction	I sell plumbing supplies. That’s a plum job, I guess!
	D	Mutation	Dried peas are my favorite snack. Dried bees – really?

**Table 2.** Examples of each type of pun as produced by Ekdom.

(2)	A	‘Feestelijk nieuws voor zanger <b>Matt Bellamy</b> . Hallo, <b>met Bellamy</b> .’
	B	Ja, ze is zwanger! Ze is in december <b>uitgerekend</b> . Dat dat nou <b>uitgerekend</b> in december is.
	C	Doen we er ook <b>handboeien</b> bij? Dan kun je geboeid zitten kijken. CD van Frank <b>Boeijen</b> .
	D	Dan moeten ze me in een <b>kist</b> de studio uit dragen. Ik laat me voorlopig niet <b>kisten</b> .
	E	Heb je gehoord hoe die plaat gaat? Is dat alleen met <b>bassen</b> ? Nee, is dat alleen met <b>Bassie</b> ?
	F	Dan heb je misschien een dodelijke erectiespin. Daar heb je dan twee uur <b>lol</b> van. Daarna is de <b>lul..</b> lol er snel vanaf.

### 2.3 The corpus

A total of 284 radio show recordings were used to build the corpus of puns. The recordings date from 15-04-2014 to 01-05-2018. Shows that aired with a different host than Ekdom himself, when he was unavailable to host, were not included in the corpus.

The annotation process of the radio shows by Gerard Ekdom resulted in a total of 937 puns, dating from 2015 to 2018, in total. After the corpus was finished, all puns were checked for the criteria to match Otake and Cutlers (2013) Japanese corpus as close as possible. All puns that were not based on a source word produced by someone were not included in the corpus. All target words (puns) that were not produced by Gerard Ekdom himself were also excluded from analysis. This yielded in a total of 816 puns that were used for analysis. The data were also checked for who produced the source word the puns were based on. A total of 627 (76.8%) source words are produced by Ekdom himself and 189 (23.2%) source words were produced by someone other than Ekdom. These other people include guests, callers and news presenters, among others. All target words, and thus puns that were used for analysis, were produced by Ekdom himself.

All puns were annotated in the corpus including a variety of information relevant for analysis. All different categories of annotated information can be found in table 3. Not all categories were necessary to answer the research questions in this study. Embedded puns and mutation puns include an extra set of information that is relevant for these specific types of puns.

**Table 3.** All information included in the corpus of puns.

(3)	Attribute	Details
<b>All puns</b>		
	Interlocutor	Who was the interlocutor the punster is interacting with?
	Pun type	What type (according to the three types of Otake and Cutler, 2013) of pun was made? (homophone, embedding or mutation)
	Source word	What is the source word?
	Target word	What is the target word?
	Who said source word?	Who produced the source word?
	Who said target word?	Who produced the target word?
	How many times SW?	How many times was the source word produced (possibly by multiple people)?
	Existing TW	Is the target word an existing word? Checked in the online dictionary VanDale
	TW before SW	Was the target word produced before the source word?
	Word type SW	What is the word type of the source word?
	Word type TW	What is the word type of the target word?
	Across word boundary	Does the target word or the source word cross a word boundary?
	Length SW	Length of the source word in syllables
	Length TW	Length of the target word in syllables
	Emphasis	Do both the source word and the target word have an identical stress pattern?
	Emphasis SW	Stress pattern of the source word (1 = stressed syllable * = unstressed syllable)
	Emphasis TW	Stress pattern of the target word
	Collocation	Is the pun part of a (mutated) collocation?
	Language SW	What is the language of the source word?
	Language TW	What is the language of the target word?
	Inflection	Is the pun conjugated?
	Preceding sentence	What sentence preceded the pun?
	Sentence with pun	What is the full sentence that includes the target word (the pun)?
	Comments	Possible particularities of or comments on the pun in question
<b>Embeddings</b>		
	Type of embedding	What is the type of embedding? Insertion or extraction
	Position overlap	In what position of the longer word does the overlap occur?
	Homophone possible	Is a homophonic pun possible? Verified via the online dictionary Van Dale

<b>Mutations</b>	
Amount of changed sounds	How many of the sounds in the target word are present not in the source word?
Amount of new sounds	How many sounds of the source word are not present in the target word?
Difference (new – old)	What is the difference between new sounds and changed sounds?
Sounds SW	Which sounds are not present in the target word?
Sounds TW	Which sounds are not present in the source word?
Position in word	In what position of the word does the mutation occur?
Homophone possible	Is a homophonic pun possible? Verified via the online dictionary Van Dale.

## 2.4 Procedure

To collect the data for the corpus of puns, the Dutch radio show Effe Ekdom was recorded. The recordings were recorded and saved as mp3 files and opened and listened to in Windows Media Player. Each recording was listened to, and all puns were located and annotated. Before this could be done, an annotation scheme was made (see table 3). This scheme was partly based on information used in the study by Otake and Cutler (2013). This was done to make comparisons with the Japanese study easily possible. Some adaptations to the annotation scheme of Otake and Cutler (2013) were necessary to include all important information of Dutch puns. The Japanese corpus included yamato and kango words, that are respectively Japanese and Chinese loanwords. Dutch does not include this type of loanwords, this category was thus not included in the Dutch corpus. A more detailed explanation of the difference between the Japanese and Dutch corpus can be found in the discussion chapter. All puns with a source word that triggered a pun were then logged in a Microsoft Excel file including all the information listed in table 3 with a time stamp (among other practicalities) to easily relocate the puns at a later time, if necessary. Each pun was categorized in one of the three pun types (homophones, embedded puns, and mutations) according to the criteria by Otake and Cutler (2013) that are listed above.

## 2.5 Data analysis

To answer the questions raised in this study, different categories of puns and information were compared to each other using statistical analyses. The question that will be answered in this paper are (1) Will the MaxSP hold for the Dutch corpus as well as for the Japanese corpus? (2) Does the Dutch corpus contain fewer homophonic puns than the Japanese corpus when comparing different categories of puns? And lastly (3) Do speakers of Dutch avoid stress

shifts from the source word to the target word when making puns? The first question is the main question answered in this study and questions 2 and 3 are sub-questions.

To answer the first research question the homophone category was compared to embedded puns and mutation puns. This was done to find out whether a homophone is the most preferred pun in Dutch as well as in Japanese. For the second question, a category of puns that changes the stress pattern was compared to one in which the puns retain the stress pattern of the source word. To answer the third, and last, research question, occurrences of pun types were compared to the Japanese study to see whether the homophone category is relatively smaller in Dutch than in Japanese. A descriptive study of all the collected data in the corpus was conducted as well and is included in the next chapter.

### 3. General Results

Much like Otake and Cutler's (2013) original study I have used non-parametric tests to analyze source words and the context they appear in. Two-tailed binomial tests were used to test whether the distribution of puns across two categories was significant. All tests were considered significant at  $p < .001$ , this was chosen to match the Japanese study as close as possible.

To give the current study more power and volume, the analyses in this section include all puns that were made by Gerard Eklom and not only the puns that were based on source words produced by someone other than Eklom himself. This is different from the study on Japanese puns. This decision was made because of the 816 puns, only 189 were based on a source word produced by somebody other than Eklom himself. Chapter 4 will give a comparison between the two different data sets: the one that will be discussed now and the one only including source words pronounced by another interlocutor.

#### 3.1 Descriptive analysis of the corpus of Dutch puns

Table 4 shows a summary of the collected data that make up the Dutch pun corpus. The table is based on the results table in the Japanese study by Otake and Cutler (2013). The table included in the current study to be able to closely replicate the results in the original Japanese study and so that comparisons can be made easily. The mean length of the source word is measured in syllables rather than moras since the Dutch language does contain moras, but they are of much less importance than they are in Japanese.

**Table 4.** Summary of the Dutch pun corpus.

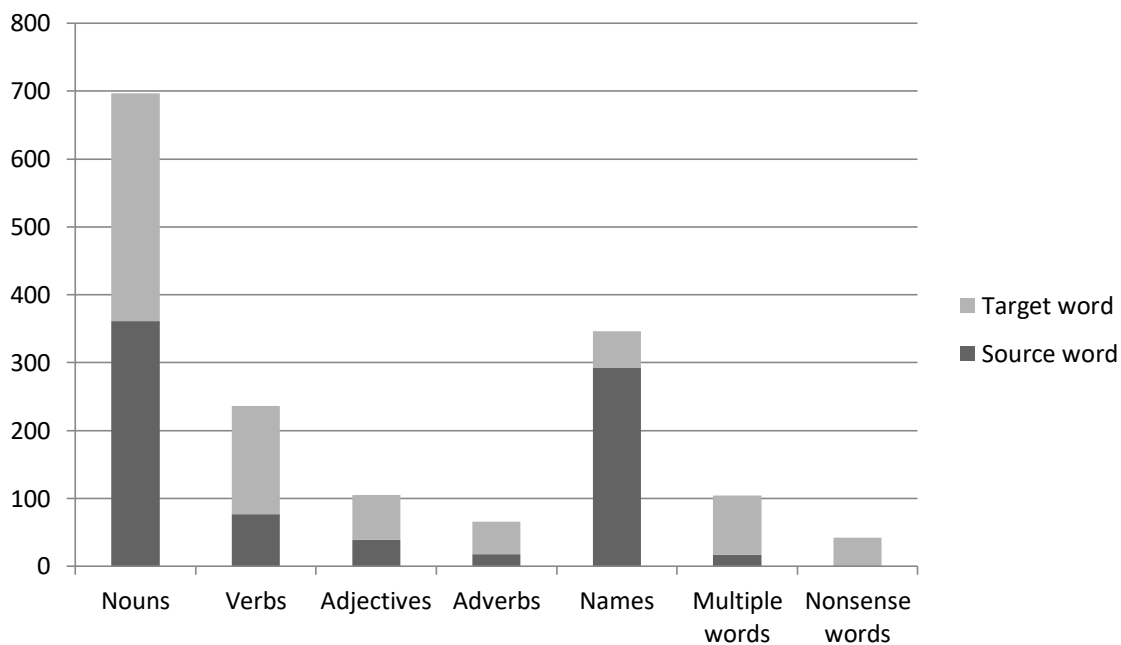
Pun type	N	% Noun		% possibility alternative pun type		% stress match source word and pun	Mean length source word (syllables)
		Source word	Target word	H	M		
Homophone	379	44.6	33.8	-	53.0	88.9	1.83
Embedding	255	46.3	49.8	32.5	-	81.2	2.10
Mutation	182	40.7	44.5	18.7	-	92.3	2.05
Total	816	44.2	41.2	-	-	87.3	1.96

### **3.2 Source word characteristics**

To replicate the Japanese study, the corpus was first analyzed for word type. Since the Japanese study resulted in a corpus with only nouns as source words chosen to be punned upon, they concluded that the choice of source words must have been meaning driven. This should also be the same for Dutch and was thus included in the analysis. The distribution of word types across all puns in the Dutch corpus for both source words and target words can be seen in figure 1. The Dutch corpus contains 816 source words, of which 361 (44.2%) are nouns. Names and proper nouns take up about 35.9 % of all source words. Other word types that appear more than once in the corpus are verbs, adjectives, adverbs, multiple word phrases, and pronouns. Abbreviations, one letter words, articles, numerals, and prepositions each occur in the corpus only once. In the Japanese study, Otake and Cutler (2013) report on the different etymology of Japanese words. They state that Japanese words can be of Japanese origin (yamato words), of Chinese origin (kango words) or be a loanword from, for example, English. Since the Dutch language does not portray this phenomenon as importantly as the Japanese language, these results are not relevant to the current study and will thus not be included in the analyses.

Of all source words, 58.3% of them were changed into another type of target word. The biggest difference is seen for names and proper nouns. Of the source words, 293 were names or proper nouns and for target words, the amount of names and proper nouns is only 53. Source word type was not changed to create the target in 29.9% of cases. In 11.6% of cases, the source was changed into a phrase of multiple words.

**Figure 1.** Distribution of the most frequent word types in the Dutch pun corpus.



### 3.3 Pun type and the MaxSP

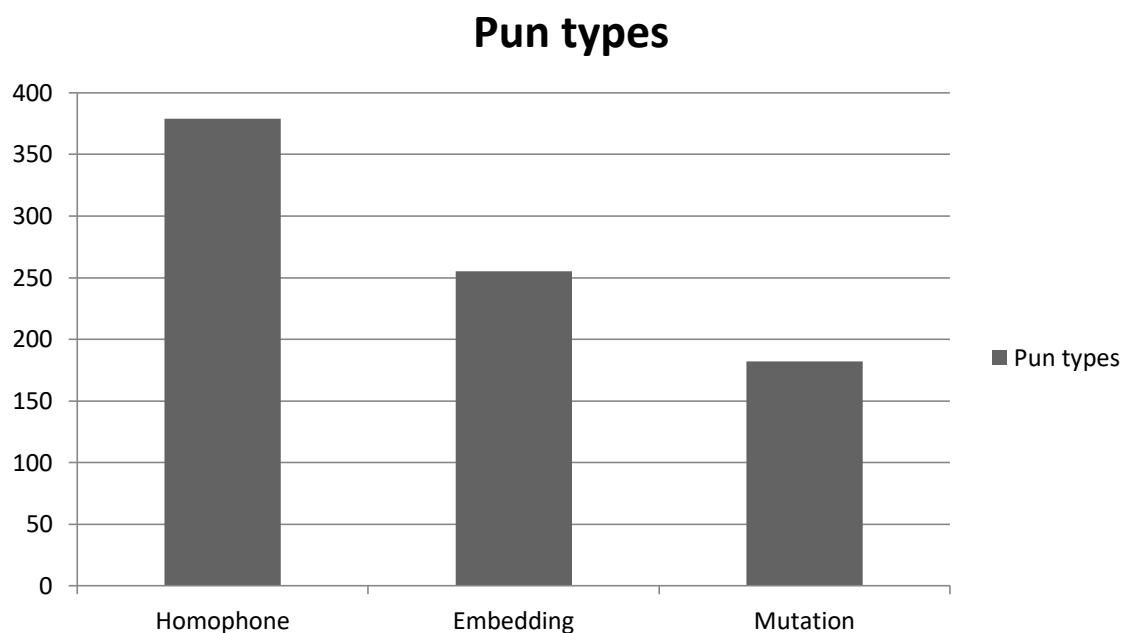
The first research question asked whether the MaxSP is also true for Dutch and was answered by comparing the three pun type categories with each other. The MaxSP states that the best and most used type of pun is a homophone (Otake and Cutler, 2013). This is because the best type of pun preserves as much of the source word the pun is based on as possible. For the Japanese study, this hypothesis turned out to be true. Of the total of 816 Dutch puns, 379 (46.4%) are homophones, 255 (31.3%) are embedded puns and 182 (22.3%) are mutations. Homophones appear in the corpus significantly more often than mutations ( $N = 182$ ,  $p < .000$ , 2-tailed) or embeddings ( $N = 255$ ,  $p < .000$ , 2-tailed). The difference between homophones and non-homophonic puns is not significant ( $p = .046$ , 2-tailed) in the Dutch corpus of spontaneous puns. Figure 2 shows the distribution of pun types across the Dutch corpus and figure 3 shows the percentages of pun types compared to the Japanese results found by Otake and Cutler.

Another way to keep the preservation of a source word is to embed it into another word. The Dutch corpus contains a total of 255 embeddings. Of those 255 embeddings 181 are insertions of a word into another, and thus preserving the source word in its entirety, and 74 are extractions. This difference is significant ( $p < .000$ , 2-tailed). The two preserving categories significantly outnumber the non-preserving category of mutation puns ( $p < .000$ , 2-tailed).

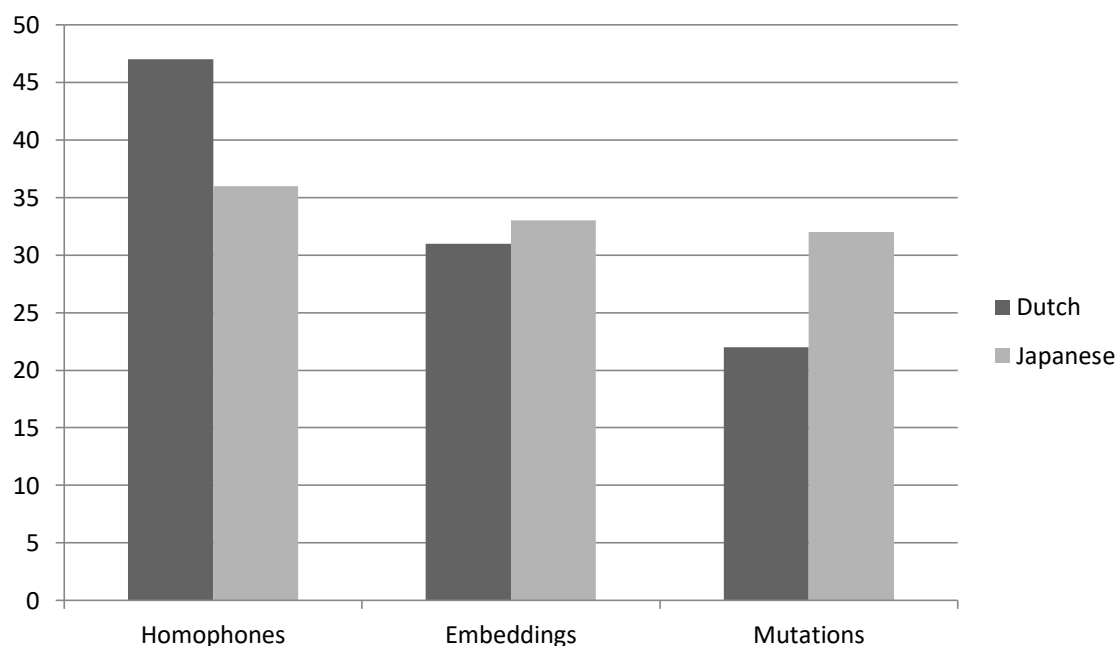


The second research question asked whether the homophone category in Dutch is smaller compared to the one in the Japanese corpus of Otake and Cutler (2013) since Dutch has a bigger phoneme repertoire. This turns out not to be the fact. The difference between the category of homophones and the others is in Dutch even bigger than it is in Japanese. A visualization of this comparison can be seen in figure 3.

**Figure 2.** Distribution of pun types in the Dutch pun corpus.



**Figure 3.** Percentages of pun types in the Dutch pun corpus compared to the Japanese corpus of Otake and Cutler (2013).



### 3.3.1 Homophones

As has already been mentioned, the MaxSP predicts the best pun type to be a source preserving homophone pun. This turned out to also hold for Dutch puns and not only for Japanese puns.

It is however not the case that no other pun type is possible for puns that are homophones. Like in Otake and Cutler's (2013) study, most homophone puns in the corpus could be a different word by mutating at most one or two sounds in the word. The pun can thus change into a different type of pun: a mutation. This mutation possibility was examined for Dutch as well, by making up words by changing (inserting, deleting or replacing phonemes) one or two sounds in the word and checking the existence of the word with an online dictionary. Of all homophones in the Dutch corpus, 201 puns can be mutated to another word with minimal changes to the word. This is, however, not a significant difference from the 178 homophones that cannot be mutated ( $p = .258$ , 2-tailed). Just as many homophones can be another type as homophones that cannot. This means that homophones were not chosen as puns because there was no other option, but rather because a homophone is a better pun and preserves more of the source word than any other type of pun.

### 3.3.3 Mutations

The MaxSP by Otake and Cutler (2013) predicts that a mutation will occur only when a source preserving option is not available. For all puns, the possibility of a homophone was registered and checked with an online dictionary. It turns out that this is not true for all cases. Of the mutations, 18.7% of puns has an option for the punster to choose a homophone. The group for which a homophone is possible is, however, a much smaller group of puns than the group of mutation puns that do not have a homophone that the punster could have chosen ( $p < .000$ , 2-tailed).

There seems to be no criterion for what type of word can be chosen to be a mutation pun. The speaker does not systematically choose words to create mutations with. There seems to be no pattern or obvious preference of a certain word length (number of syllables). The words chosen to be mutated are on average 2.05 syllables long and the distribution of word types is similar to the total of all puns in the corpus. The distribution of word types for the three different pun types can be seen in figure 3 and figure 4. The change of a source word to a nonsense word occurs more often in embeddings and mutations than in homophones. It seems to be a preference for homophones to be existing words.

### 3.3.2 Embeddings

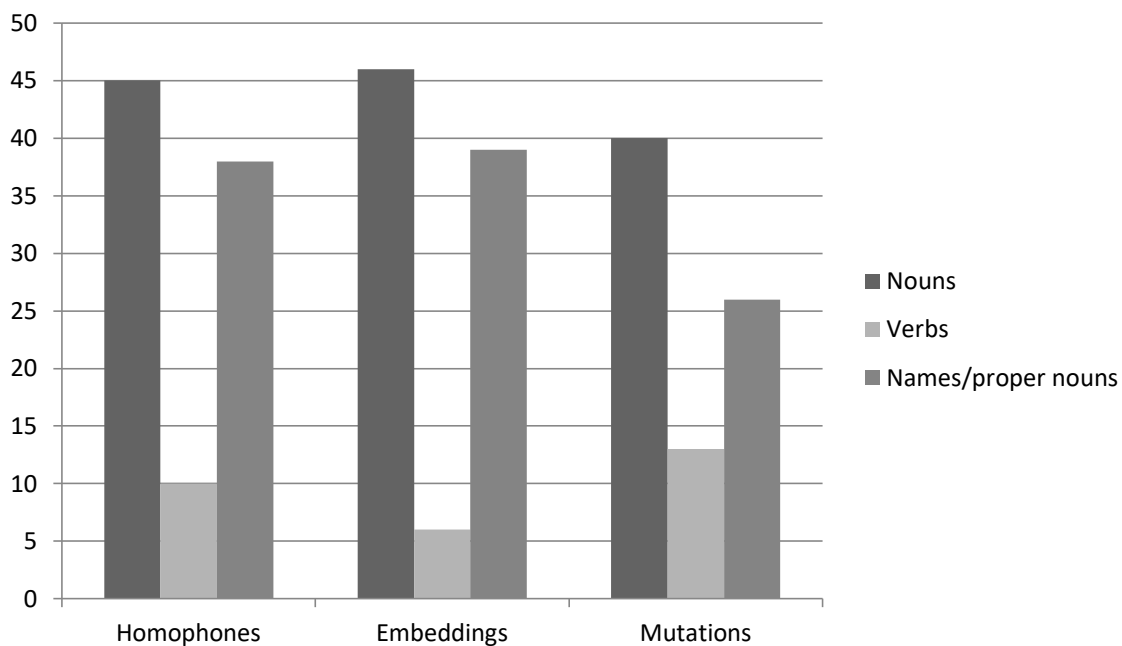
Of the embeddings, only 83 embedded puns (32.5%) have a homophone available that could have been chosen by the punster, and 172 embeddings do not have a homophone to choose ( $p < .000$ , 2-tailed).

Embedded puns are subdivided into two different types of embeddings: insertions and extractions. An insertion happens when a source word is embedded in a longer carrier word while extractions are target words that are extracted from a longer source word. As has already been mentioned, a significant difference has been found between the two different types of embedded puns. Most embedded puns are insertions of a source word into a longer word. This difference does not show in the analysis of word length in syllables. Only a small difference in word length is shown. The source words for embedded puns have a mean length of 2.10 (SD = .97) syllables while the target words are longer ( $M = 2.63$ ;  $SD = 1.30$ ). The difference is, however, negligible.

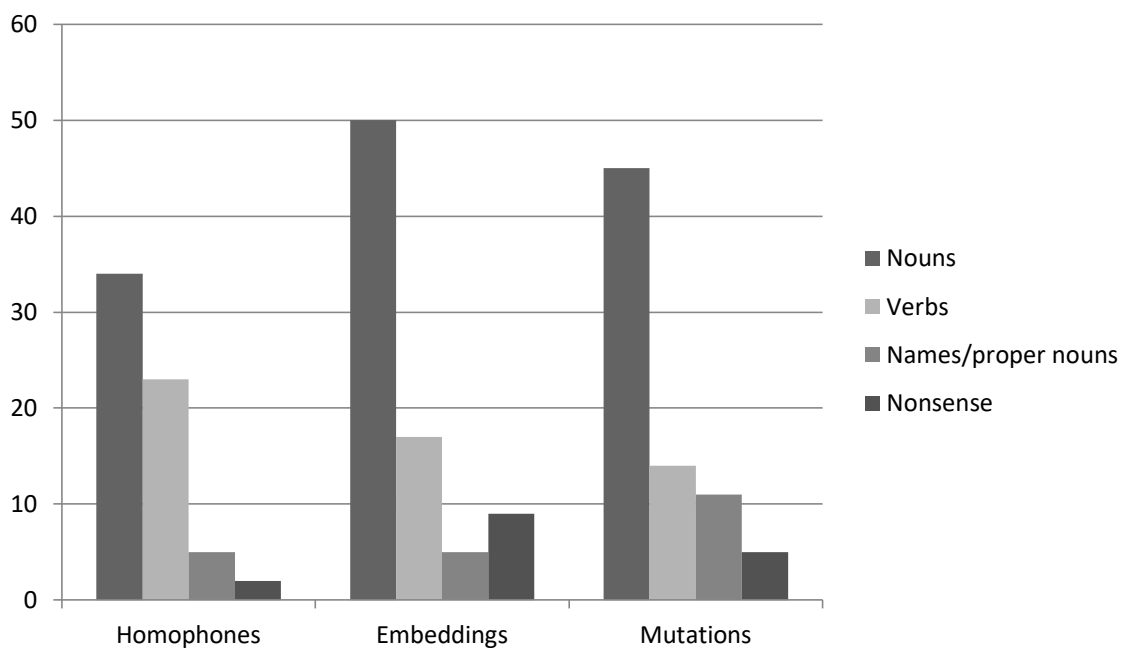
I have mentioned in the introduction that an extraction of a word from another word is more likely to come from a word of more than one syllable since that is easier to understand for the listener in the word recognition process. The data show, however, that it is possible to extract a word from a word of only one syllable. However, this occurred only 10 times, which

is a lot less than an extraction from a word of more than one syllable. An extraction from a source word longer than one syllable occurred 64 times. This is a significant difference.

**Figure 3.** Percentages of the three most occurring source word types per pun type.



**Figure 4.** Percentages of the three most occurring target word types per pun type plus nonsense words.

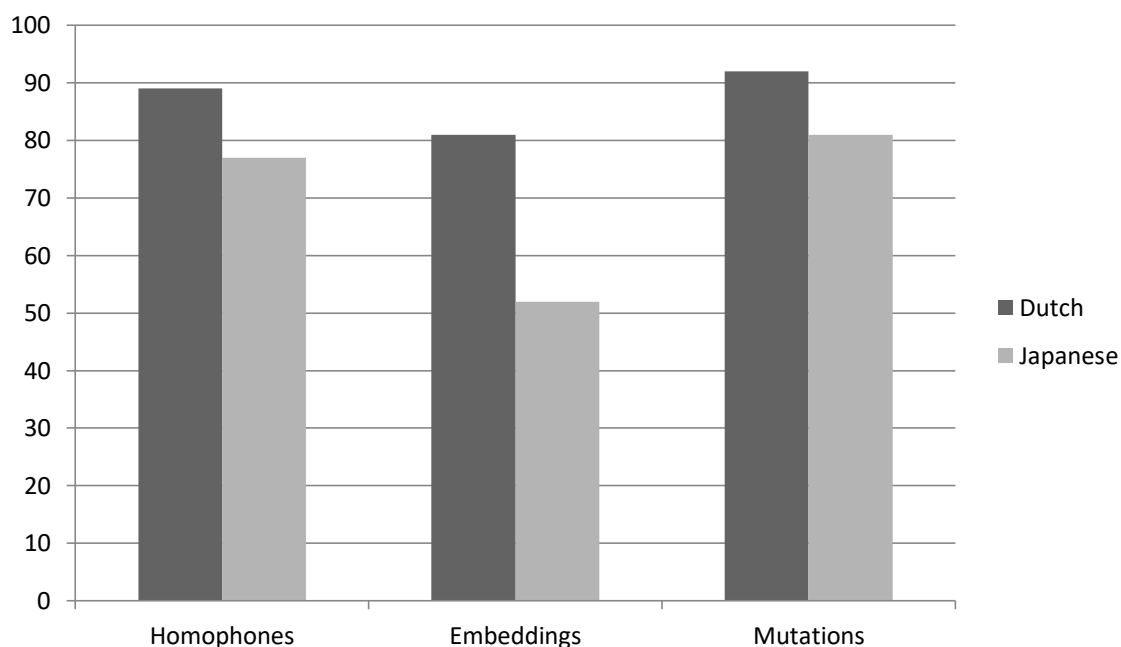


### 3.4 Source word preservation and stress

The current study focused partly on what properties of a source word are possible to be changed in a target word when crafting a pun. One of these properties is stress. The third research question asked whether stress shifts between source words and target words are disfavored in Dutch. Otake and Cutler (2013) have mentioned that in English stress shifts are not a favorable option. In English, it is best to preserve stress, according to Otake and Cutler. In Japanese, this appears to be less important. The Dutch data show a pattern similar to that predicted for English. Of all puns, 12.7% show a stress shift while 87.3% does not. Similar to the Japanese study on homophones 88.9% of the Dutch data preserves stress while Otake and Cutler (2013) report 77% stress preservation. For mutations, this is 92.3% in Dutch and 81% in Japanese. A big difference with the Japanese study was observed for embedded puns. Of all Dutch embedded puns, 81.2% preserves the stress pattern of the source word while in Japanese this is only 52%. Figure 5 visualizes these differences between the two studies.

These results are in line with the Principle of Maximal Source Preservation of Otake and Cutler (2013). The MaxSP also predicts how many sounds can be or are preferred to be changed in a mutated source word. The Japanese data showed that in most mutation cases, very little of the source word is changed to form the target word. Changing only one phoneme is the preferred option for mutation puns in Japanese. This is also true for the Dutch data. Of all mutations in 62.1% one phoneme was changed, for 17.6% two phonemes were changed and for 4.9% of mutation puns, three phonemes were changed. Changing four or more phonemes only occurred twice in the entire corpus. The difference between one and two or more phoneme changes is significant ( $p < .000$ , 2-tailed).

**Figure 5.** Percentage stress preservation per pun type Dutch vs Japanese.



### 3.5 Position in word

To stay true to the study of Otake and Cutler (2013) the corpus was analyzed to try to replicate the Japanese results on the position of alterations in embedded puns and mutation puns. Otake and Cutler concluded, based on analysis of their corpus, that in Japanese word-initial positions are most likely to be altered in a mutation pun. For embedded puns, word-initial overlap is most likely to be the basis of a Japanese embedding. The data in the Dutch corpus shows no replication of these results. A replication was not expected since initial mutations can affect word recognition in Dutch. Most overlap in embeddings in the Dutch corpus happens in the final position of words. This happens in 49.8% of all embeddings while 44.3% of all embeddings show a word-initial overlap. The difference between these positions is, however, not significant ( $p = .524$ , 2-tailed) and can be seen as no difference at all. Dutch embedded puns do not have a preference for the position of overlap.

Puns that show initial mutation also linked to the neighborhood model of lexical access. This theory states that when a word like ‘lamp’ is heard, a word like ‘camp’ is activated as well as ‘lamp’. This makes puns that change the first sound and/or the onset understandable as puns. For mutation puns, the Dutch corpus shows 71 (39%) word-initial mutation and 44 (24.2%) word-final mutations. This difference is, different to the Japanese study, not significant ( $p = .015$ ).

## 4. Differences in source word production

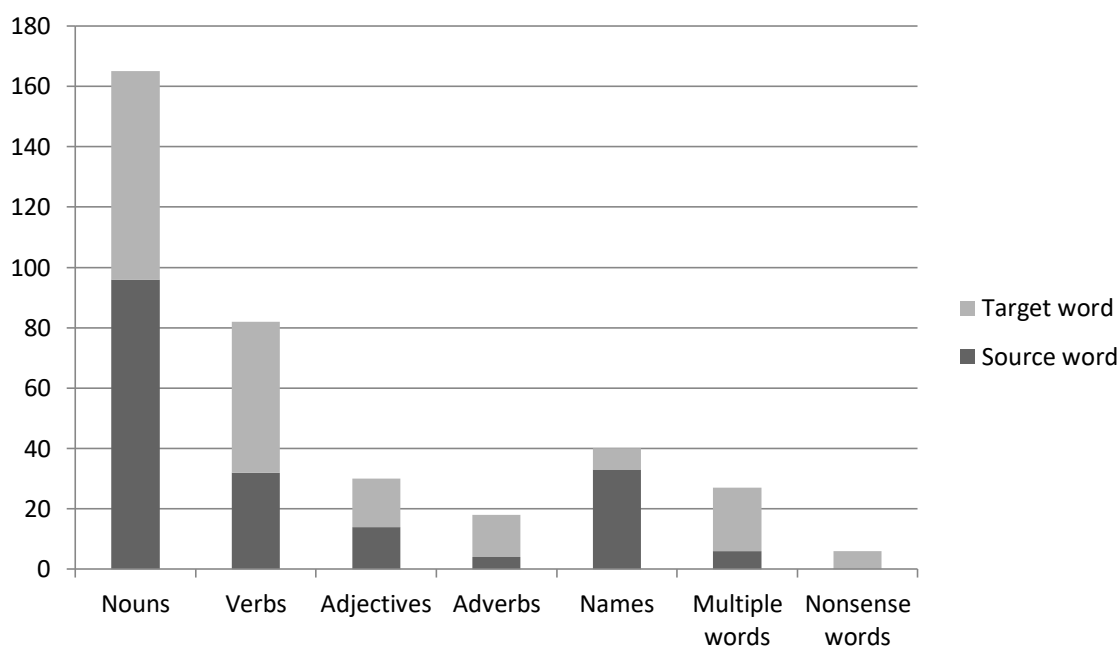
All source words in Otake and Cutler (2013) were produced by somebody other than the radio presenter that their research was based on. In all previous analyses of the current study on Dutch, all source words were included to give more power and volume to the research. Since Otake and Cutler (2013) only used puns based on source words produced by other people, this section will discuss results for the puns that were based on a source word produced by someone other than Gerard Etkind. All analyses from the previous section will be run again for the 189 puns that have a source word produced by somebody other than Etkind. A comparison with the results from the previous section will be made.

### 4.1 Source word characteristics

Figure 6 shows the distribution of word types across the source words for this part of the data. The figure is the same as the one in the previous section. This data set consists of 189 source words, of which 96 (50.8%) are nouns. Of all source words, 17.5% are names or proper nouns. Other word types that appear are verbs, adjectives, adverbs, multiple word phrases, and pronouns, just like in the data in the previous section.

For 58.7% of the source, words the type of word was changed for the target word. Source words that are names and proper nouns show the biggest word type difference between source and target word. Of the source words, 33 (17.5%) were names or proper nouns. This is very different for target words, of which the amount of names and proper nouns is only 7 (3.7%). In 29.6% of puns, the word type was not changed to create the target word. Sometimes, a source word was changed into a phrase containing multiple words. This happened in 11.6% of cases, which is exactly the same proportion as for the full corpus that included all puns in the previous section. The rest of the numbers only show small differences with the results shown in the previous section but are almost identical.

**Figure 6.** Distribution of the most frequent word types for puns based on source words produced by someone other than Gerard Ekdöm.



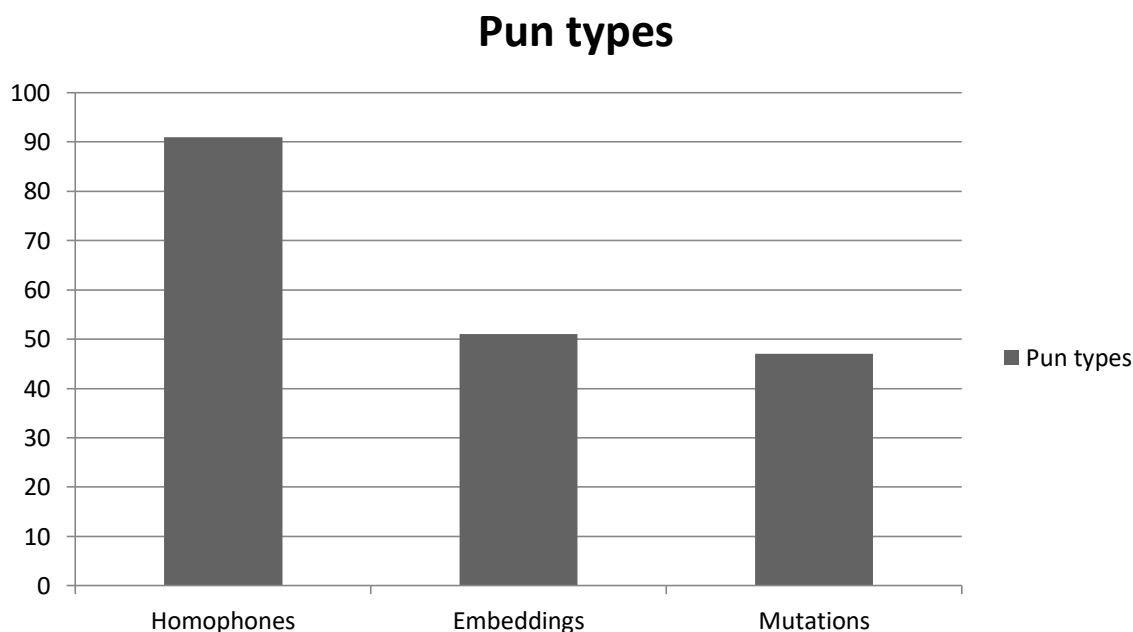
## 4.2 Pun type and the MaxSP

In the data set containing Dutch puns based on a source word the punster did not produce 91 puns (48.1%) are homophonic puns, 51 puns (27.0%) are embedded puns and 47 puns (24.9%) are mutations. The category of homophonic puns ( $N = 91$ ) is significantly bigger than mutations ( $N = 51$ ,  $p < .001$ , 2-tailed) or embeddings ( $N = 47$ ,  $p < .000$ , 2-tailed). The difference between the categories of homophonic and non-homophonic puns is in this data set not significant ( $p = .663$ , 2-tailed). In this data set, the difference between embedded puns and mutation puns is much smaller than the difference between the two for the entire corpus. The distribution can be seen in figure 7.

Keeping as much of the source word the same as the target word can also be done by embedding a source word into a larger carrier word or extracting a target word from a longer source word. This part of the Dutch corpus contains a total of 51 embeddings. Of those 51 embeddings, 34 are insertions of a word into another and 17 are extractions from a source word. This difference is not significant ( $p = .024$ , 2-tailed), this could, however, be due to the small amount of data used in this analysis. The data set in the previous chapter was much larger. When held to different standards of significance, the difference is, in fact, significant ( $p < .05$ ). Homophones and embeddings, which preserve the source word, significantly outnumber the non-preserving category: mutations ( $p < .000$ , 2-tailed).



**Figure 7.** Distribution of pun types for puns based on a source word not produced by the punster.



#### 4.2.1 Homophones

Like in the previous section, most homophone puns in the corpus could be changed into a different word by mutating at most one or two sounds in the word. Of all the homophones in this data set, 47 homophonic puns can be mutated to be another word by changing only one or two phonemes. The difference between these homophones and the 44 homophones that cannot be mutated is not significant ( $p = .834$ , 2-tailed).

#### 4.2.2 Mutations

Like in the previous section, some of the mutation puns do have a homophone that could have been chosen by the speaker. This was true for 6.4% of mutation puns. But most mutation puns do not have a homophone to choose from. The group that does not have a homophone possibility is significantly bigger than the group of mutations that does ( $p < .000$ , 2-tailed). This is also a much bigger difference than the one seen in the previous section where 18.7% of puns had a homophone to choose.

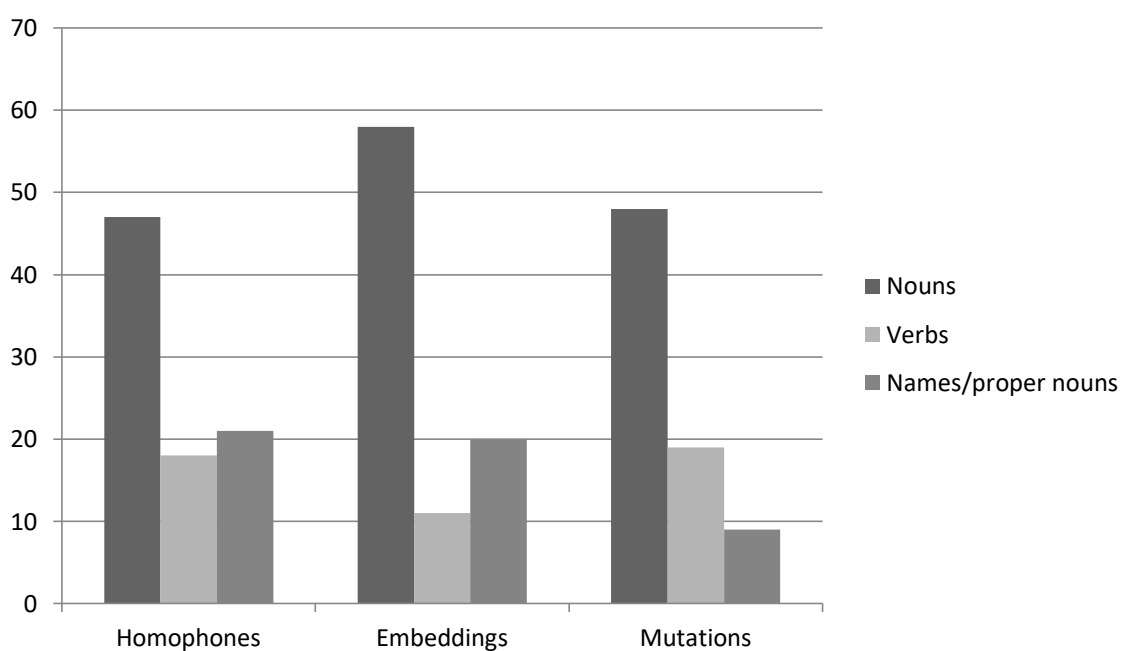
The source words that the punster chooses to mutate into a different word are on average 2.15 (SD = .807) syllables long while this was 2.05 syllables for the entire corpus. The distribution of word types for all pun types can be seen in figure 8 and figure 9. Just like in the entire corpus the change of a source word to a nonsense word occurs in this data set more often in embeddings and mutations than in the category of homophones.

### 4.2.3 Embeddings

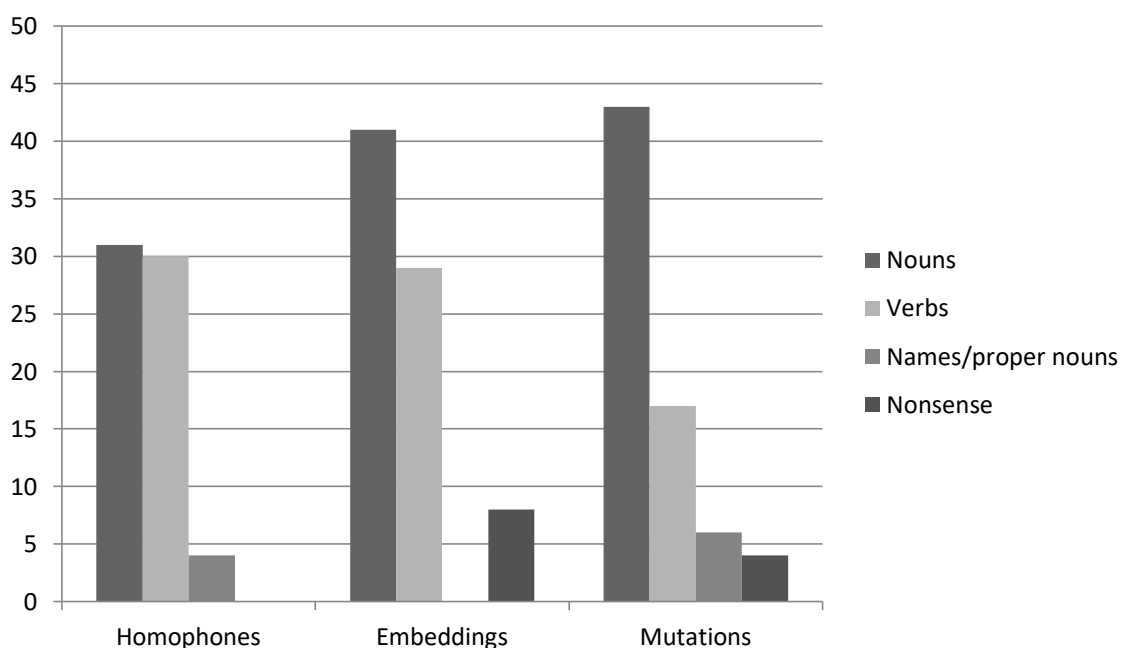
Only 19 (37.3%) of the embedded puns have a possible homophone that could have been chosen and 32 embedded puns do not have a possible homophone ( $p < .092$ , 2-tailed). This difference is, likely due to the smaller sample than in the previous section, not significant.

Insertions of source words into carrier words occur more often than extractions. This result did not show in the analysis of word length in syllables in the previous chapter and did also not show in this current subset of the data. A small difference in word length can be observed. The mean length of source words is 2.14 (SD = 1.00) syllables, while the target words are a little bit longer (M = 2.53; SD = 1.19). The difference is, however just like in the previous chapter, negligible.

**Figure 8.** Percentages of the three most occurring source word types per pun type



**Figure 9.** Percentages of the three most occurring target word types per pun type plus nonsense words



### 4.3 Source word preservation and stress

The current data set shows a pattern like the data from the previous chapter. Of all the homophones 87.9% of this part of the Dutch data preserve stress while Otake and Cutler (2013) report that 77% of Japanese homophones preserve stress. For mutations, this is 91.5%, 92.3% in the previous chapter and 81% in Japanese. For embedded puns, 81.2% of all Dutch embedded puns and 80.4% of this subset of the data preserve the stress pattern of the source word, while this is 52% for Japanese puns.

This subset of the data is in line with the MaxSp of Otake and Cutler (2013). Their theory also entails how many sounds are preferably mutated in a mutation pun. The Japanese data and the data discussed in the previous chapter showed that in most cases there are minimal changes to the source word to create the pun, which means that changing only one phoneme is the most preferred option. The fewer sounds that are changed, the better the pun will be, according to the MaxSP. This turns out to hold for the current data set of puns based on a source word produced by someone that is not Ekdom. Of the mutations in this data set, for 66.0% of puns only one phoneme was changed, for 19.1% two phonemes were changed and for 2.1% of mutation puns, three or more phonemes were changed. This is a significant difference ( $p < .000$ , 2-tailed).

#### 4.4 Position in word

In Japanese, according to Otake and Cutler (2013), alterations in mutations are most likely to occur in word-initial positions. For embedded puns, word-initial overlap is most likely to be the basis of a Japanese embedding. The data in the entire Dutch corpus, which was reported in the previous section, shows no replication of these results. The current data set does not replicate these results either, which was expected since a replication did not happen in the previous section either. Most overlap in the data analyzed in the previous chapter happens in the final position of words. In this subset of the corpus, there is no preference for the position in the word at all. All three categories show an almost identical percentage. For mutations, 36.2% show a word-initial mutation, 29.8% of mutations happen in the middle of the word and 27.7% of all mutations have a word-final position. For embeddings, the results are quite different. Of the embedded puns 52.9% show overlap in initial positions, 41.2% show final overlap while only 3.9% show overlap in medial positions of the word. This result is, however, not surprising due to the nature of embedded puns. It can be concluded that, different to Japanese, Dutch mutation puns and embedded puns do not have a preference for position of overlap.

#### 4.5 Difference in data sets

Table 5 shows all (very small) differences in the percentages of the analyzed information between all puns and puns based on a source word produced by someone other than the punster himself. The percentage of nouns is bigger in the second data set, but still much lower than the Japanese percentage of nouns. There are thus almost no differences between the two compared data sets and this will not be discussed further. All results can be safely taken together to draw conclusions from without making any difference to the conclusions that will be drawn. Only the results of section 3 will be discussed in the following chapter.

**Table 5.** Differences between the data in chapter 3 and chapter 4.

Pun type	% Pun types		% Stress preservation		% Noun	
	All puns	Different SW production	All puns	Different SW production	All puns	Different SW production
Homophone	46.4	48.1	88.9	87.9	44.6	47.3
Embedding	31.3	27.0	81.2	80.4	46.3	58.8
Mutation	22.3	24.9	92.3	91.5	40.7	48.9
Total	100	100	87.3	86.8	44.2	50.8

## 5. Discussion and conclusion

This study focused on the act of spontaneous punning and its relationship to the way the Dutch language is made up, produced and recognized. The study was mostly based on a study that was done by Otake and Cutler (2013) on puns in the Japanese language. This Japanese study was partly replicated but there were some differences made in the methodology and additional research questions were added to match the Dutch language. The first and the last research questions could be answered using the Dutch data set alone while the second question was based on a comparison between Japanese and Dutch and the other two. These research questions that have been answered in the current study are: (1) Will the MaxSP be true for Dutch as well as for Japanese? (2) Does the Dutch corpus contain fewer homophones than what has been found in the Japanese study? And lastly: (3) Do speakers of Dutch avoid stress shifts when crafting puns? The current study used a corpus analysis of spontaneously produced Dutch puns based on a study by Otake and Cutler (2013) on Japanese. The collecting of puns resulted in a total of 816 puns to be used for analysis. All information that was included in the corpus has been mentioned in this paper so the results and methodology of this study would be repeatable.

The corpus was made up of puns produced by a Dutch radio show host and contains hundreds of spontaneously produced puns based on a clear source word. The corpus was annotated with different information about each pun based mostly on an annotation scheme from the Japanese study by Otake and Cutler (2013). The most important categorization in the corpus was the types of puns that Otake and Cutler introduced: homophone puns, mutation puns, and embedded puns. Detailed information on the corpus can be found in the method section and table 2.

### 5.1 The MaxSP

It has been hypothesized for the first, and main, research question, which is: Will the Principle of Maximal Source Preservation be true for Dutch as well as for Japanese?, that the MaxSP as introduced and proven to be true for Japanese puns will also be true for puns made in Dutch. That means that is expected that homophones will be the biggest category of puns since the Principle of Maximal Source Preservation means that a pun that remains as similar to the source word as is possible. More changes or deviation is disfavored. This is what Otake and Cutler (2013) state in their paper on Japanese puns. The best and most occurring pun, according to Otake and Cutler (2013), should thus be a homophone. The first research

question of this study has been answered by comparing the category of homophonic puns to the other two categories: mutation puns and embeddings. The MaxSP states that the best pun is a homophone and this should thus be the biggest category of puns in the corpus. It was hypothesized that this should be true for Dutch as well as for Japanese. It turns out that homophones make up 46.4% of the corpus, while 31.3% of all puns are embeddings and 22.3% are mutation puns. The differences between homophones and the other categories are significant. It can be concluded that the hypothesis that a homophone is the best pun (the MaxSP) is true for Dutch. The result found for Japanese by Otake and Cutler (2013) has been replicated for Dutch. No difference was found when the distribution of pun types of both languages was compared.

The Principle of Maximal Source Preservation turns out to not only be true for Japanese, but also for Dutch. This means that it could also hold across even more languages or even be a universal across all languages. More research regarding other languages is necessary to prove this speculation.

Otake and Cutler (2013) further explore the question whether the MaxSP is only used for the quality of puns or if it is also used by a speaker during the spontaneous process of punning. It turns out that the MaxSP is an important factor when punning spontaneously in Japanese. The current study showed that this is also a true fact for Dutch spontaneous punning and it can be concluded that the MaxSP plays an important role in the production process of puns. Even when they are produced spontaneously and when they have not been written or prepared in advance.

Since Otake and Cutler (2013) made a distinction between source words and included only source words produced by someone else in their study, a comparison between puns based on a source word produced by someone else and puns based on a source word produced by Gerard Eklom himself was made. This resulted in no difference and all puns can be taken together for drawing conclusions without it making any difference to the results.

## **5.2 Pun types**

We have already seen that the most occurring pun type in the entire corpus is the category of homophone puns. This has been explained by the Principle of Maximal Source Preservation of Otake and Cutler (2013). The best pun is one that preserves as much of the source word as possible. In a homophone pun, the source word and the target word stay entirely the same, although it does in rare occasions differ in stress pattern. The second most preserving pun type is the type of embedded puns and especially insertion embeddings. Insertion means that

the punster embeds the source word into a longer carrier word to create the pun. This turned out to be the second biggest category in the corpus. The MaxSP would predict the category of insertion embeddings to be bigger than the one of extractions. Insertions turn out to occur more often than extractions in both the Dutch study and the Japanese study. The smallest category is that of mutated puns. Mutation puns are the category that preserves the least of the source word because in this category the punster does not only add or subtract from the source word, he or she changes one or more sounds of the word. The MaxSP can explain this distribution of pun types across the corpus very well.

However, not all puns seem to fit clearly into one category. During the annotation process, it became evident that some puns can be classified in two different pun type categories. Some examples from the Dutch corpus are listed below. For these puns, the most salient type, based on the context of the words, was chosen as the classification.

The first example comes from the embedding category: source word ‘fan’, target word ‘fanster’. The word ‘fanster’ can possibly be a newly crafted word for a female fan of a band or artist, but can also be a homophone of ‘venster’, which means ‘window’, since in the variation of Dutch our punster speaks the /f/ and /v/ phonemes sound almost identical. I have chosen to categorize this case as an embedding of the source word ‘fan’ into ‘fanster’.

Another example was also found in the embedding category: source word ‘geit’ (goat) or ‘gegeten’ (eaten), target word: ‘gegeiten’. In this case, both the word ‘geit’ and ‘gegeten’ can be seen as the source word, because they are both produced in the conversation in question. If ‘geit’ is the source word, the classification would be an embedding of ‘geit’ into ‘gegeiten’. If ‘gegeten’ is seen as the source word, it would be a mutation where the phoneme /e/ is switched for /ei/. Since the conversation mainly contained the subject goats I have chosen to classify this case as an embedding.

The last example can be either a homophone or an embedding: source word ‘nageltjes’ (small nails), target word ‘genageld’ (nailed). Due to inflection, a verb is made from a noun by embedding ‘nagel’ into it. The question whether inflection played a role was added to the corpus for these types of cases. The pun can then be classified as a homophone, since both the source word and the target word are based on ‘nagel’, but have undergone an inflection process for, in this case, a diminutive and a verb. Since there are so few of these cases, relative to the size of the corpus, it seems unlikely to have made a big difference to the results of this study and can thus be seen as negligible.

### **5.3 Phonetic and prosodic makeup**

Not only the type of pun and the MaxSP plays a role in the act of punning, the way a language is made up, both phonetically and prosodically, can also possibly affect the way puns are made in a language. The second and third research questions focused on this possibility. The second research question, and first sub-question of this paper, which was: Does the Dutch corpus contain fewer homophones than what has been found in the Japanese study?, was answered by comparing the difference between homophones in this study and other categories to those in the Japanese study by Otake and Cutler. It has been hypothesized that homophones will appear less in the Dutch data than in Japanese since the Dutch phoneme repertoire is larger than that of Japanese and this may affect how many homophones the language holds and how they are used and treated in the language in use.

It turns out that the homophone categories in both languages are very close to each other. In Japanese 36% of all puns are homophones while in Dutch 47% of all puns are homophones. It has been hypothesized that in Dutch, the relative category would be smaller than that of Japanese. This hypothesis has been proven wrong by the results of the corpus analysis. The category of homophones in Dutch turns out to be even bigger than the one in Japanese. Phonetic makeup does not seem to play as big of a role in punning as has been widely believed.

The fact that the homophone category in Dutch is just as big as the one in Japanese is not compatible with the hypothesis I have made in the first chapter that Dutch would have relatively fewer homophone puns since it has a bigger phoneme repertoire than Japanese. Several sources state that for a language with fewer phonemes it is a necessity to include more homophones in the vocabulary because with fewer phonemes, fewer unique words can be formed. It should, therefore, be so that a language with more phonemes in its repertoire should include less homophones. This happens to not be the case for the act of punning. For puns and wordplay, it turns out that preservation of the source word is more important than the phonetic makeup of the language. Puns can, however, reflect the vocabulary of a language in general. Otake and Cutler (2013) compared their corpus of puns to the Japanese language in general. They state that around 36% of all words in Japanese are homophones and that this number is the same as what they found in their corpus. Unfortunately, this information on Dutch was not readily available to me. A comparison between the number of homophones in the Dutch language in general with the number of homophones in this corpus of Dutch puns could help us understand the relationship between language structure and the act of punning to an even greater extent.



Another interesting finding in the results of the current study is the fact that in Japanese, most, if not all, puns contain nouns, while in the Dutch corpus only less than half of all source words are nouns. Otake and Cutler (2013) found almost all source words that were chosen by the Japanese punster to be nouns and a similar pattern was found for the target words. For Dutch, the pattern is very different. Otake and Cutler state that the bias for nouns comes from the fact that the process of choosing a word to pun on is meaning driven. A noun also has an independent meaning that does not necessarily draw on the context it appears in. They also state that the answers the guests in the Japanese radio show almost always contain a noun. The difference with the Dutch corpus could be attributed to either the difference in radio show content and design or personal preference of the punster. I do, however, speculate that this could be due to differences in the language. Otake and Cutler state that the phonology of Japanese yields candidate words with many activated neighbors that are often homophones. This could be different to Dutch because of the already mentioned differences in phonetic makeup. The Dutch phoneme repertoire is larger than that of Japanese and while this did not affect homophone frequency, it could have affected the distribution across and the choice of source words.

Dutch homophones can also be something different than just nouns and can change their word type across the different meanings. Cutler and Van Donselaar (2001) give an example of this. ‘Voornaam’ can in Dutch be both a noun and an adverb. The noun means ‘first name’ while the adverb means ‘distinguished’. An example for English was given by Small, Simon, and Goldberg (1988). The word ‘convert’ can mean two different things for the noun and the verb. The noun of such words is usually the result of what is done in the verb. A condition for these examples in both Dutch and English is that the lexical stress shifts between the words. For example, the noun for convert is stressed like CONvert while the verb is stressed as conVERT.

Besides the difference in nouns, an interesting difference in use of stress shifts was found between Dutch and Japanese. The third research question was based on differences in the use of stress. This question asked whether stress shifts in puns are disfavored in Dutch, was answered by comparing the number of puns that shift the stress pattern to the number of puns that do not. It has been hypothesized that, since in English stress shifts are not preferred in puns and English and Dutch are quite similar, stress shifts in puns are not the favored option in Dutch. The data in the corpus do reflect this hypothesis. Of all puns, 87.3% does not change the stress pattern from the source word the pun is based on. Only 12.7% does shift the

stress when making the pun. When the stress of the different categories is compared to the results in Japanese, Dutch homophones and mutations show a similar pattern to Japanese.

Embeddings, however, show a very different pattern. In Japanese only, 52% of puns preserve the stress pattern of the source word, while in Dutch this is 81.2%. As has been hypothesized, stress shifts in Dutch puns are disfavored. This is compatible with studies on stress. It has been found by Cutler and Clifton (1984) that prosodic information can be used to facilitate word recognition. Misstressed words are more difficult to identify than words that are correctly stressed. Cutler, Cooper, and Wales (2002) have stated that a differently stressed word can remove possible word candidates of the acoustic signal a listener receives. This is a possible reason for the disfavoring of stress shifts in puns. When a word candidate is rejected, the pun is possibly understood very slowly or not at all. Better puns are thus puns that preserve the stress pattern of the source word. With the MaxSP in mind, this is not at all surprising, since preservation plays a key role in the Principle of Maximal Source preservation.

#### **5.4 Puns and word recognition and production processes**

Puns, we have seen, are in fact, like McGraw and Warren (2010) have hypothesized, based on two (sometimes contradicting) meanings that elicit a humorous effect. Puns come in pairs of two words: source words and target words. Whether the source word and target word differ minimally in form or meaning, or even both, they come in, sometimes contradicting, pairs.

The making of these puns is based on normal word recognition and production processes. This has already been established by Otake and Cutler (2013) but has been proven again in the current study. When punning, a punster first needs to choose a source word he or she will base his or her pun on. This word needs to be recognized and will activate other words in its neighborhood. McQueen and Cutler (1992) explain this as a model that depends on how many other words a word that is heard resembles at any point. These words will then get activated in the listener's mind. They give the word 'lice' as an example. The recognition of 'lice' will be affected by the fact that 'lie', 'eye' and 'ice' are embedded inside the word 'lice'. This neighborhood activation is what makes the act of punning possible in the first place. This model can also play a role in the recognition of words that are embedded in other words.

After the choice of the source word, the punster needs to choose a target word from all activated associates. This target word needs to be accessed and produced by the speaker. When the to be produced word is selected, semantically related words will also be activated in

a speakers mental lexicon, according to Levelt (2000). Much like what happens in a word recognition process. After the selection of the lemma, syntax is added to the utterance that will be produced. And after the syntactic properties of the word are accessed, phonological properties of the word need to be accessed and then produced by the speaker. When dealing with synonyms, both phonological codes will get activated. For homophones, an important type of pun in this study, it is the case that both words can be pronounced with slightly different durations (Warner, Jongman, Sereno and Kems 2003). The activation of two different phonological codes means that when producing a homophone, the second meaning needs to be accessed from a different place than the first meaning. The words that are seemingly the same are stored as two different phonological codes. After activation of the phonological code, the pun can be produced by the speaker and the listener will need to recognize and understand the pun they have just heard. The listener then goes through the word recognition process, the same way the speaker did when choosing a source word.

The Neighborhood Model explains both the choice of a source word by the activation of related words and why puns work and why they are understood and elicit laughter from the listener. When a listener hears the pun, the same recognition process happens again. The target word will activate the source word in the mental lexicon. This works, even when there is no clear source word that was produced. Some puns are plays on existing words that have not been produced before making the pun. These types of puns were not included in the current study, but can be researched as well. Especially the reception of these types of puns will be interesting to take a look at since the source word needs to be accessed from the lexicon and not from short-term memory. That is because it has not been said in the conversation the pun appears in. Interesting insights can be derived from studying puns without a source word in the context the pun appears in.

## **5.5 Limitations and recommendations for further research**

The conclusions made in this paper are subject to limitations. A first limitation comes from the sources of the source words chosen to be punned upon. In the results section, I have stated that there is no difference between puns based on a source word that has been produced by Gerard Ekdorf himself or by someone else. Otake and Cutler (2013) did not include any puns based on source words that were produced by the speaker himself. A side note to this statement is, however, that a lot of source words produced by Ekdorf did come from somewhere other than his speech alone. Most source words produced by Ekdorf were either read from written news items, read from the internet or read from comments of listeners that

were received by text. It is, however, very difficult to know where each source word actually came from, especially because there were no visuals to the data used in this paper. Had this information been more readily available, a more reliable split between the source word origins could have been made. A better knowledge of the origin of source words may result in significant differences.

Another limitation of this study is the fact that only one person and only one language were studied. While the fact that only one speaker was included in the corpus did give the study more reliability, the analysis cannot lead to much universals. For all conclusions drawn in this paper, it could be personal preference of the speaker and not a universal for the language or puns in general. More research on different or multiple speakers could resolve this issue and confirm the statements made about these data. This study also only includes a comparison between Japanese and Dutch. Research on different languages can possibly confirm the theory that the best puns are generally puns that preserve as much of the source word as possible. This confirmation can lead to universals for all languages.

Another point of attention and a side note to the results is the use of identical puns in the data. These data do, in fact, include a lot of puns that occur more than once in the corpus. This might have an effect on the results. However, it seems entirely justified to not erase identical puns from the data since Dynel (2010) stated that a lot of jokes and puns are based on witticisms people have already heard or used before that are stored in their memory and mental lexicon. This storing of words in the mental lexicon is also the case for normal word production processes and I expect that this won't change anything about the results and conclusions drawn from the data in this study.

## **5.6 Conclusions**

This study has explored the use of puns in spontaneous speech in the Dutch language and has compared the Dutch data to a study on Japanese by Otake and Cutler (2013). The first research question was based on the Japanese study and asked whether their Principle of Maximal Source Preservation is true for Dutch spontaneous puns as well as it was for Japanese. The principle states that the best pun is a pun that stays as close to the source word as possible. This would mean that the homophone category in a corpus should be the biggest. This has been hypothesized for Dutch and turned out to be true. The MaxSP is true for Dutch as well as for Japanese.

The second research question dealt with the fact that Dutch and Japanese are languages that have a different phonetic make up. The difference is in the phoneme

repertoires of both languages that differ in size. A bigger phoneme repertoire could mean that a language has fewer homophones in its vocabulary. It has thus been hypothesized that for Dutch since the phoneme repertoire is bigger, the category of homophones should be smaller in relation to the one in Japanese. The Dutch corpus was compared to the Japanese corpus and the two homophone categories turned out to be the same size in relation to the other categories. Phonetic makeup does thus not play as big a role in pun production as preservation of the source word.

The third, and last, research question asked whether stress shifts are not preferred in Dutch. Since stress can play a big role in word recognition and understanding it was hypothesized that stress shifts in puns would not appear as much as stress retaining puns in the Dutch corpus. This hypothesis was proven to be true for Dutch. Stress does play a role in the act of spontaneous punning.

Despite the limitations, I have discussed and the questions that this study undoubtedly raised, this study did give a good insight in the language processes in action and their relationship with the phonological and prosodic makeup of the Dutch language. In general, it can be concluded that the act of punning draws on normal word recognition and word production processes that belong to the language the puns are made in. The Principle of Maximal Source Preservation is an important factor in this process. The best pun is, in fact, a homophone.

## 6. Bibliography

- Bowers, J. S., Davis, C. J., & Hanley, D. A. (2005). Automatic semantic activation of embedded words: Is there a “hat” in “that”? *Journal of Memory and Language*, 52(1), 131-143.
- Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(6), 1430.
- Cole, R. A., Jakimik, J., & Cooper, W. E. (1980). Segmenting speech into words. *The Journal of the Acoustical Society of America*, 67(4), 1323-1332.
- Costa, A., Cutler, A., & Sebastián-Gallés, N. (1998). Effects of phoneme repertoire. *Perception & psychophysics*, 60(6), 1022-1031.
- Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and speech*, 45(3), 207-228.
- Cutler, A., & Clifton Jr, C. (1984). The use of prosodic information in word recognition. In *Attention and performance X: Control of language processes* (pp. 183-196). Erlbaum.
- Cutler, A., & Van Donselaar, W. (2001). Voornaam is not (really) a homophone: Lexical prosody and lexical access in Dutch. *Language and speech*, 44(2), 171-195.
- Cutler, A., Norris, D., & Sebastián-Gallés, N. (2004). Phonemic repertoire and similarity within the vocabulary. In *Eighth International Conference on Spoken Language Processing*.
- Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes*, 16(5-6), 507-534.
- Dynel, M. (2010). How do puns bear relevance?. *Relevance studies in Poland*, 3, 105-124.
- Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, 84(3), 474-496.

- Hahn, U., & Bailey, T. M. (2005). What makes words sound similar?. *Cognition*, 97(3), 227-267.
- Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, 92(1-2), 101-144.
- Isel, F., & Bacri, N. (1999). Spoken-word recognition: The access to embedded words. *Brain and Language*, 68(1-2), 61-67.
- Kawahara, S., & Shinohara, K. (2011). Phonetic and psycholinguistic prominences in pun formation: Experimental evidence for positional faithfulness. *Japanese/Korean linguistics*, 18, 177-188.
- Levelt, W. J. (1999). Models of word production. *Trends in cognitive sciences*, 3(6), 223-232.
- Levelt, W. J. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences*, 98(23), 13464-13471.
- Leyden, K. V., & van Heuven, V. J. J. P. (1996). Lexical stress and spoken word recognition: Dutch vs. English.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25(1-2), 71-102.
- McGraw, A. P., & Warren, C. (2010). Benign violations: Making immoral behavior funny. *Psychological science*, 21(8), 1141-1149.
- McQueen, J. M., & Cutler, A. (1992). Words within words: Lexical statistics and lexical access.
- McQueen, J. M., Norris, D., & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(3), 621.
- Miller, T., Hempelmann, C., & Gurevych, I. (2017). SemEval-2017 Task 7: Detection and interpretation of English puns. In *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)* (pp. 58-68).
- Norris, D., McQueen, J. M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology*, 34(3), 191-243.

- Otake, T., & Cutler, A. (2001). Recognition of (almost) spoken words: Evidence from word play in Japanese. In: *Seventh European Conference on Speech Communication and Technology* (pp. 465-468)
- Otake, T., & Cutler, A. (2013). Lexical selection in action: Evidence from spontaneous punning. *Language and speech*, 56(4), 555-573.
- Partington, A. S. (2009). A linguistic account of wordplay: The lexical grammar of punning. *Journal of Pragmatics*, 41(9), 1794-1809.
- Ritchie, G. (2005). Computational mechanisms for pun generation. In *Proceedings of the Tenth European Workshop on Natural Language Generation (ENLG-05)*.
- Rodd, J., Gaskell, G., & Marslen-Wilson, W. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language*, 46(2), 245-266.
- Sandra, D., Frisson, S., & Daems, F. (1999). Why simple verb forms can be so difficult to spell: The influence of homophone frequency and distance in Dutch. *Brain and language*, 68(1-2), 277-283.
- Shinohara, K., & Kawahara, S. (2010). Syllable intrusion in Japanese puns, dajare. In *Proceedings of the 10th meeting of Japan Cognitive Linguistic Association*.
- Small, L. H., Simon, S. D., & Goldberg, J. S. (1988). Lexical stress and lexical access: Homographs versus nonhomographs. *Perception & Psychophysics*, 44(3), 272-280.
- Van Ooijen, B. (1996). Vowel mutability and lexical selection in English: Evidence from a word reconstruction task. *Memory & Cognition*, 24(5), 573-583.
- Van Mulken, M., Van Enschoot-van Dijk, R., & Hoeken, H. (2005). Puns, relevance and appreciation in advertisements. *Journal of pragmatics*, 37(5), 707-721.
- Vroomen, J., & De Gelder, B. (1997). Activation of embedded words in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 23(3), 710.
- Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of phonetics*, 32(2), 251-276.