sempre:
Society for Education, Music and Psychology Research

Article

Psychology of Music
2020, Vol. 48(4) 547–563

DOI: 10.1177/0305735618812199
journals.sagepub.com/home/pom
SAGE

Individual differences in the reminiscence bump of very long-term memory for popular songs in old age: A non-linear mixed model approach

Daniel Zimprich

Abstract

Recognition of popular songs from one's past shows a phenomenon that is known as a "reminiscence bump" from autobiographical memory research, namely, the increased recognition performance of songs from one's youth and early adulthood. As a first goal of the present study, a non-linear functional relation between popular song recognition performance and song-specific age of an individual was examined. As a second goal, individual differences in recognition performance curves were taken into account by including random effects. The third goal was to explain individual differences by including predictor variables. The sample comprised 90 participants aged 70 to 75 years. Participants listened to excerpts of 51 songs of the German charts from the years 1945 to 1995. Results show that the average bump performance was 75%, that the bump was located at about 17 years of age, and that the inflection point was located at 31 years. Individual differences could be explained by the number of correctly recognized songs, musical taste during one's youth, and the frequency and preference of listing to bump songs. To conclude, an individual differences approach based on a non-linear function relation has been found to be a promising way to understand why popular songs from one's youth are remembered better.

Keywords

non-linear mixed model, popular songs, recognition, reminiscence bump, very long-term memory

Me, I blame the old songs,
Filling hearts with rebellion and romance

– Joe Jackson "Old Songs"

Department of Developmental Psychology, Ulm University, Germany

Corresponding author:
Daniel Zimprich, Department of Developmental Psychology, Institute of Psychology and Education, Ulm University, Albert-Einstein-Allee 47, Ulm, 89081, Germany.
Email: daniel.zimprich@uni-ulm.de

There are songs we never forget. Even after years, hearing a small excerpt of one of these songs may suffice to remember its melody and lyrics, the interpreter of the song, things that happened during listening to that song, people we were together with at that time, etc. Despite this seeming effortlessness, music recognition is a complex procedure. Minimally, recognizing a song involves a comparison between sensory input and a representation of some of the properties of the musical piece (Hébert & Peretz, 1997). The phenomenon of recognizing music heard a long time ago is one manifestation of what may be called "very long-term memory" (Platz, Kopiez, Hasselhorn, & Wolf, 2015; Schulkind, Hennis, & Rubin, 1999). Other exemplars of very long-term memory may involve recalling public events (Janssen, Murre, & Meeter, 2008), the contents of books (Larsen, 1996), prose (Rubin, 1977), or sports events (Rubin, Rahhal, & Poon, 1998). What makes very long-term memories intriguing is that they cover a long time-span, that they stem from the real world and that they, oftentimes, have a personal meaning (Neisser, 1978).

Compared to memory research conducted in the laboratory (e.g., Kurtz & Zimprich, 2013; Zimprich, Rast, & Martin, 2008), the examination of very long-term memories offers less experimental control. One reason is that the encoding phase is not part of the experiment, which is why it cannot completely be ensured that, e.g., all participants have heard a specific song before. This lack of experimental control over the encoding phase affects recognition performance, because a participant cannot recognize a song never heard. The experimenter can, however, increase the likelihood that *most* participants have heard *most* songs by selecting popular songs. If, in turn, songs are too well known, the task of recognizing them becomes too easy. Thus, one has to balance the difficulty of the recognition task by selecting popular, but not overly famous songs (Holbrook & Schindler, 1989; Platz et al., 2015).

Moreover, while in laboratory studies, the time between encoding and recall is filled with other, independent tasks (e.g., Zimprich, Martin, et al., 2008), the years that have passed since having heard a song are filled with idiosyncratic activities. Some of these activities may involve re-listening to a song. In addition to these individual "rehearsals," public media contribute to repeated "stimulus presentation." A song may, for example, continue to receive airplay or may be used in a commercial. Notwithstanding, the investigation of long-term memories is promising in that it may come closer to how memory functions in everyday life (Janssen et al., 2008, Rubin et al., 1998; Schulkind et al., 1999).

Performance in a long-term memory task involving songs could be measured by summing up the number of recognized songs – as it would be in a laboratory task (e.g., Zimprich, Rast, et al., 2008). If the songs differ in the year they were popular, there is an alternative way to quantify memory performance. One could then measure recognition performance across the age of the songs, i.e., the time passed since the songs were popular. Equivalently, instead of focusing on the age of the songs, one could use the song-specific age of the participants, i.e., the age of participants when a song was popular. Thus, from a long-term memory perspective, it is not the total number of recognized songs that is of interest, but rather how recognized songs are related to the age of participants. In other words, in research of very long-term memory for popular songs a *recognition performance curve* across age can be examined.

From the perspective of a "classic" forgetting curve (Zimprich & Kurtz, 2013), one would expect that the more distant in time from now a song was popular, the less likely it is recognized. However, previous studies have shown that recognized songs[1] mainly stem from one's juvenile years (Bartlett & Snelus, 1980; Platz et al., 2015; Schulkind et al., 1999; Zimprich & Wolf, 2016a). That is, songs that were popular during one's adolescence are recognized better. This phenomenon of an increased recognition performance of songs from one's youth relative to other life periods is similar to a phenomenon pertinent in autobiographical memory, the reminiscence bump (Rubin & Schulkind, 1997; Wolf & Zimprich, 2015; Zimprich & Wolf, 2016b).

The reminiscence bump describes the finding that when adults are asked to recall autobiographical memories, a disproportional high number of autobiographical events from adolescence is reported (Rubin et al., 1998). Thus, unlike classic forgetting curves, it is not the temporal distance of an autobiographical event that governs its recallability. Rather, it appears to be a certain life period, namely, a person's adolescence, which strongly affects the recallability of autobiographical events. Similarly, if a song was popular during an individual's youth, the likelihood of it being recognized is greater (Schulkind et al., 1999; Zimprich & Wolf, 2016a).

Several suggestions have been made as to why adolescence is special for the memory for popular songs. First, musical taste develops during adolescence (Holbrook & Schindler, 1989). One reason is that many popular songs address typical problems adolescents face (love, sexuality, rebellion, etc.). Schwartz and Fouts (2003) found adolescents prefer listening to lyrics that focus on the developmental issues they are dealing with. Therefore, there appears to be a correspondence between an adolescent's life and the songs' lyrics. This correspondence helps in developing a person's self-identity – a developmental task important during adolescence (e.g., Erikson, 1950). The way problems of adolescence are addressed in songs may function as an exemplar of how to think, feel, and act (Larson, 1995). In that sense, during adolescence music may serve as a means to regulate emotions. In line with this, Lonsdale and North (2011) found younger participants to be more likely to listen to music with the goal of regulating their emotions.

In addition, music "is one of the most common modes of self-expression among young people" (Rentfrow, 2012, p. 410). Different types of music characterize different teenage subcultures. In order to belong to a subculture, adaptation to the music genre dominant in that subculture is, at least in part, mandatory. As such, a social identity formation process occurring during adolescence is accompanied by specific music listening habits, which affect one's musical taste. This may be especially important if a subculture is relatively distant from mainstream culture and, thus, also from mainstream musical taste. Hence, individual music preferences may represent a person's values, beliefs, and lifestyle choices (Rentfrow, 2012). In that sense, the music an adolescent listens to serves as an "identity badge" that provides others with information about themselves (North & Hargreaves, 2004).

To summarize, adolescence and young adulthood are special for the long-term memory for music because musical taste and listening habits develop during this period and accompany both *personal and social identity* development. Consequently, the music listened to during this life period may be recognized best because it was or is significant for one's identity.[2] Therefore, one would expect that recognition performance is best for songs stemming from a participant's adolescence (Schulkind et al., 1999; Zimprich & Wolf, 2016a).

In a study in which songs are used as stimuli, recognition performance is expected to depend on musical taste because those with a less mainstream musical taste are less likely to have heard popular songs. As Schindler and Holbrook (2003) postulated, musical taste during adolescence influences lifelong preferences, implying that it predicts recognition performance. Moreover, we expected that the extent to which popular songs from an individual's adolescence are still listened to and whether participants enjoy listening to them now also contributes to recognition performance (Conway, 2005). Finally, the total number of recognized songs should also be a powerful predictor of a recognition performance curve, because it determines the area under the recognition performance curve. However, it does not contain information about how recognized songs are distributed across age – participants with the same number of recognized songs could thus exhibit different recognition performance curves (Schellenberg & Habashi, 2015).

These variables and other variables of interest could be included as predictors of a recognition performance curve in a regression-like statistical approach. In the present study, we suggest an approach based on a combined power and exponential function in order to model individual differences in the relation of recognition performance with the song-specific age. Much like in a forgetting function, the functional relation between recognition performance and the time passed since a song was popular is modeled.

Goals of the present study

We followed three goals with the intention of extending previous findings on a very long-term memory of music.

(1) What is the functional relation between song-specific age (of participants) and song recognition performance? We examined whether a combined power-exponential function can adequately describe the trajectory of recognized songs across song-specific ages of participants. Note that, previously, the lifespan distribution of memory for popular songs has been analyzed in a descriptive manner by providing, e.g., histograms (e.g., Schulkind et al., 1999). In the present study, we wanted to push the analysis level a step further by investigating whether the song recognition performance curve can be captured by a non-linear function.

(2) Do individuals differ reliably with respect to their music recognition performance curves? This question was addressed by augmenting the power-exponential function with random effects. As is typical for mixed effects models (Hedeker & Gibbons, 2006), our approach relies on the assumption that individual recognition performance curves follow the same functional form as the average curve. Individual differences are then differences in the parameters governing the non-linear function, not differences in the type of the function (Meredith & Tisak, 1990). Note that these individual differences have not been addressed in most previous research on memory for popular songs (but see Zimprich & Wolf, 2016a, 2016b).[3]

(3) To what extent can various predictors account for individual differences in very long-term memory of popular songs? As outlined above, we expected participants' musical taste during adolescence, music listening frequency, and musical preferences to affect music recognition performance curves. In addition, the total number of recognized songs was assumed to influence recognition performance curves.

Method

Sample

The sample of the present study was composed of N=90 older persons. On average, these persons were 72.5 years old (SD 1.52 years), with the youngest and oldest participants being 70 and 75 years old, respectively. Of the 90 older persons, 50 participants (56%) were women (see Table 1). The sample had an educational background typical for the general population of older persons in Germany, with 55 participants (61%) reporting to have finished a 5-year secondary school (equal to 9 years of schooling), 21 participants (23%) who finished a 6-year secondary school (equal to 10 years of schooling), and 14 participants who graduated from high school (equal to 13 years of schooling).

Material

We selected 51 songs that were ranking between positions nine and 11 of the annual German charts, that is, based on their chart performance over the course of the whole year, from the

Table 1. Descriptive statistics of the sample.

Variable	Mean	Standard deviation	Minimum	Maximum
Age	72.51	1.53	70	75
Sex[a]	0.56	0.50	0	1
Musical taste during adolescence[b]	0.74	0.44	0	1
Current musical taste[b]	0.49	0.50	0	1
Number of songs correctly recognized	25.32	8.31	4	46
Frequency of listening to bump songs[c]	4.23	1.44	1	7
Frequency of listening to non-bump songs[c]	2.55	1.19	1	6
ΔFreq (listening frequency of bump over non-bump songs)[d]	1.68	0.87	−1.31	3.94
Preference of bump songs[e]	5.44	1.33	2	7
Preference of non-bump songs[e]	3.32	1.29	1	6
ΔPref (preference of bump over non-bump songs)[f]	2.12	0.93	−1.25	4.46

[a]Dummy-coded as 0 = male, 1 = female.
[b]Dummy-coded as 0 = jazz/classical/rock/pop, 1 = German pop.
[c]On a scale ranging from 1 = never to 7 = very often.
[d]Calculated as the individual mean frequency rating of bump songs minus the individual mean frequency rating of non-bump songs.
[e]On a scale ranging from 1 = not at all to 7 = very much.
[f]Calculated as the individual mean preference rating of bump songs minus the individual mean preference rating of non-bump songs.

years 1945 to 1995 (one song per year). Most of the songs belonged to the rock/pop or German pop music genre. More detailed information about the songs is presented in the Appendix. From each song, an audio excerpt of 30 seconds was created. Importantly, the excerpts neither included the chorus nor lyrics mentioning the title.

Procedure

Participants were tested individually with a notebook computer used for stimulus presentation. After having given their informed consent, participants entered demographic data and answered questions enquiring into their musical taste and their music listening habits. Regarding musical taste, participants were asked to describe their musical taste during adolescence. Possible genres were “German pop,” “rock/pop,” “jazz,” and “classical music.” In addition, participants were asked to describe their current musical taste with the same four genres. Both variables were dummy-coded with “German pop” (taste = 1), which represented the most frequent musical taste both during adolescence and now, versus “jazz/classical/rock/pop” combined (taste = 0).

Subsequently, participants were instructed how the music memory task would proceed. Afterwards, participants were presented with the auditory excerpt of the first song. The order of the 51 songs was randomly assigned for each participant. After the excerpt had ended, participants were prompted with the question of whether the song sounded familiar. If yes, they were asked for the title and the artist of the song. If they were able to name either one, the stimulus song was considered as recognized. Participants received feedback of whether they recognized a song. Afterwards, participants were asked to rate on two 7-point Likert-type scales how much they like to listen to the song (1 = *not at all*, 7 = *very much*) and how frequently they listen to the song (1 = *never*, 7 = *very often*). This procedure was repeated for the remaining songs.

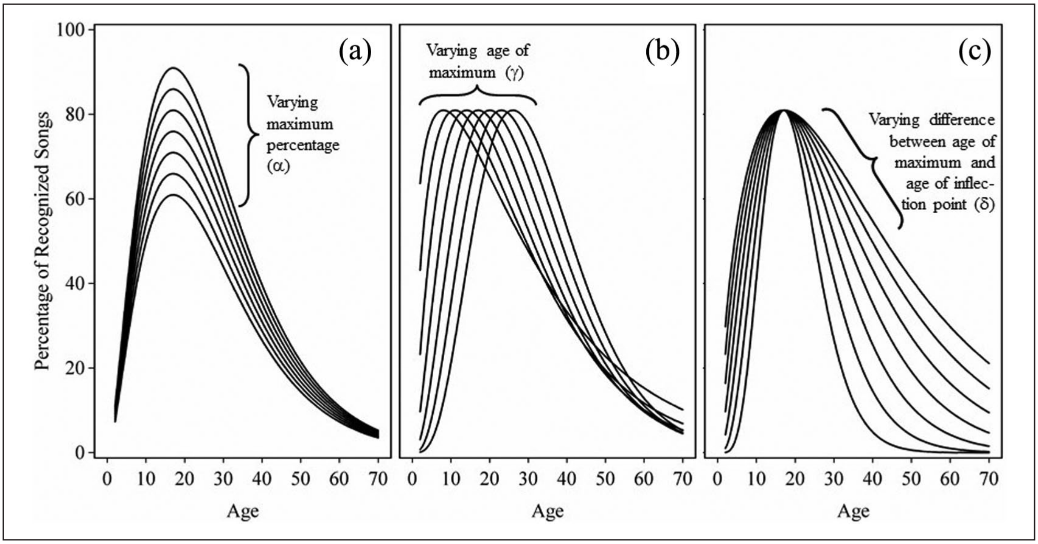

Figure 1. Modeling the combined power-exponential function: meaning of the parameters α, γ, and δ. The function is plotted for different values of the maximum percentage α (a, left panel), different values of the age where the maximum occurs γ (b, middle panel), and different values for the age difference δ between the age of the maximum and the age of the inflection point (c, right panel).

From the year a song was popular, we calculated the song-specific age of the participants. Next, song-specific ages were gathered into 5-year age bins (1–5 years, 6–10 years, etc.). The primary outcome variable of the present study was the percentage of songs correctly recognized in each age bin from that age bin. If, for example, a participant recognized four songs from the age bin of 6–10 years, the percentage of recognized songs in that age bin is $4/5 \times 100 = 80$.

Modeling approach

The relation between age bins and recognition performance was modeled using a combination of a power and an exponential function given as

$$y_{it} = \alpha \exp\left(\frac{\gamma(\gamma - t)}{\delta^2}\right)\left(\frac{t}{\gamma}\right)^{\frac{\gamma^2}{\delta^2}} \tag{1}$$

where y_{it} is the percentage of songs that were correctly recognized by participant i falling into age bin t, parameter α is the maximum value of y, that is, α=max(y), γ is the value of t where the maximum occurs, that is, $\gamma = t_{max}$, and δ is the difference between the value of t where the maximum occurs and the value of t of the inflection point, that is, $\delta = t_{inf} - t_{max}$. Parameter α, thus, captures the "height" of the reminiscence bump, parameter γ captures the "location" of the reminiscence bump, that is, at which age it occurs, and parameter δ describes the falling of the function beyond the maximum. Alternatively, from a "reminiscence bump" point of view, δ describes how pronounced the bump is – for small (large) values of δ, the bump is more (less) pronounced. Figure 1 shows how the three parameters affect recognition performance curves. In the left panel (a), the effect of different values of α is shown. Clearly, α determines the maximum of the curve or the "height" of the bump. The middle panel (b) of Figure 1 shows the effect of different values of γ. As can be seen, γ shifts the location of α, that is, it governs the age at which the bump occurs. Finally, the right panel (c) of Figure 1 demonstrates the effect of varying δ. Noticeably, δ determines how recognition performance is distributed across age, that is, how pronounced the bump is, leading to either a more "leptokurtic" or a more "platykurtic" shape of the curve.

In order to account for the dependency and to capture individual differences in the reminiscence bump, random effects for the three parameters α, γ, and δ were introduced. After including random effects denoted by Latin letters, the function becomes

$$y_{it} = (\alpha + a_i)\exp\left(\frac{(\gamma + g_i)(\gamma + g_i - t)}{(\delta + d_i)^2}\right)\left(\frac{t}{\gamma + g_i}\right)^{\frac{(\gamma+g_i)^2}{(\delta+d_i)^2}} \tag{2}$$

Under standard assumptions about random effects (zero mean, normality), these random effects can be estimated. If one assumes that the three random effects are distributed multivariately normal, the variance-covariance matrix of α, γ, and δ is

$$\begin{bmatrix} a_i \\ g_i \\ d_i \end{bmatrix} \sim \mathrm{MVN}(\mathbf{0},\Sigma) = \mathrm{MVN}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_a^2 & & \text{sym.} \\ \sigma_{ag} & \sigma_g^2 & \\ \sigma_{ad} & \sigma_{gd} & \sigma_d^2 \end{bmatrix}\right), \tag{3}$$

where MVN denotes multivariate normality, σ_a^2, σ_g^2, and σ_g^2 are the variances of the random effects, while σ_{ag}, σ_{ad}, and σ_{gd} capture their covariances.

Equations (2) and (3) represent a model that belongs to a general class of estimable models (Meredith & Tisak, 1990). However, the model in Equation (2) is non-linear, which requires non-standard parameter estimation (Davidian & Giltinan, 1995). Typically, numerical approximations are needed to maximize the likelihood (Molenberghs & Verbeke, 2005). All models were estimated using SAS NLMIXED (SAS Institute Inc., 2008), employing adaptive Gaussian quadrature. As measures of effect size, we report Cohen's d and R^2 values for both levels of analysis, that is, for the within-person (level 1) and the between-person (level 2) level (Snijders & Bosker, 1994).

Results

Table 1 contains descriptive data. As Table 1 shows, most participants reported their adolescent musical taste to have been "German pop" (74%), whereas the current musical taste was, in about equal parts, "classical/jazz/pop/rock" and "German pop." The number of songs recognized was 25.32, on average, showing that participants recognized about half of the songs. Although those with "German pop" as their adolescent musical taste recognized slightly more songs, this difference was not significant, $t(88) = 0.57$, $P > 0.55$, and the effect size was small ($d = 0.14$). Similarly, although those with "German pop" as their current musical taste recognized somewhat fewer songs, this difference was also not significant, $t(88) = -1.21$, $P > 0.23$) and corresponded to a small effect ($d = 0.18$). Participants reported listening to "bump songs", i.e., songs from their "bump age" (i.e., a song-specific age between 10 and 25 years) more frequently than to "non-bump songs." This difference was statistically significant, $t(89)=8.43$, $P < 0.01$, and represented a large effect ($d = 1.28$). Analogously, the reported preference of listening to "bump songs" was larger than that for "non-bump songs." The difference in preference was also statistically significant, $t(89) = 9.58$, $P < 0.01$, and large ($d = 1.51$).

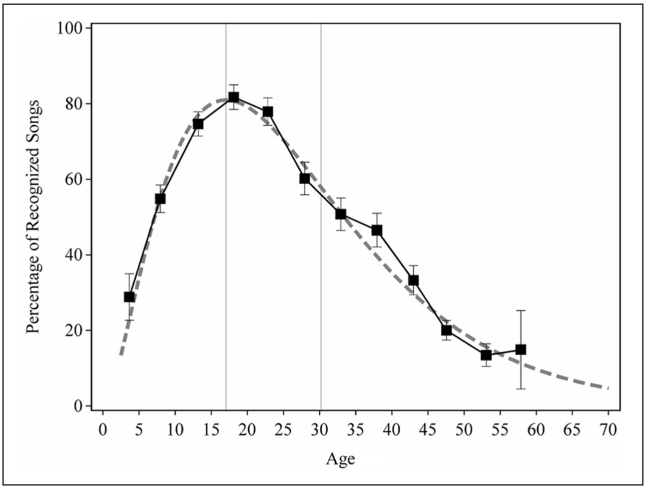

Figure 2. Percentage of recognized songs across different song-specific ages (black squares), combined with their 84% confidence interval. The broken grey line represents the best-fitting combined power-exponential function describing the data (based on model 1).

We combined the frequency variables of the bump and non-bump songs by subtracting the latter from the former. The resulting variable ΔFreq mirrors the amount of which songs from the bump age are listened to more frequently than songs from the non-bump age. As Table 1 shows, on average the listening frequency of bump songs was 1.68 higher than that of non-bump songs. However, the negative minimum value (−1.31) shows that some participants reported the listening frequency of non-bump songs to be higher than that of bump songs. Analogously, we combined the two preference variables. The resulting variable ΔPref reflects how much bump songs are preferred over non-bump songs. As Table 1 shows, the preference of bump songs was 2.12 higher than that of non-bump songs. As before, the negative minimum value (−1.25) indicates that some participants preferred non-bump songs.

Figure 2 shows the average popular song recognition performance in relation to the song-specific age of participants (gathered into 5-year age bins). There is a reminiscence bump between the song-specific age of approximately 10 and 25 years. Recognition performance rises to a maximum of about 80% at the age of approximately 17 or 19 years and decreases thereafter. Figure 2 also depicts the best fitting combined power-exponential function, which closely resembles the observed recognition performance curve.

Combined power-exponential mixed models

In a first power-exponential model (model 0), the fixed effects of α, γ, and δ were estimated.[4] Table 2 shows the results of parameter estimation. Parameter α was estimated as 75.27, implying that, on average, maximum recognition performance was 75%. The parameter estimate of γ was 16.78, indicating that the maximum recognition performance α of 75% occurred around age 17 years. The estimate of δ was 14.24, that is, the inflection point of the function describing recognition performance was 14.24 years after γ, that is, at an age of 31 years.[5] The fit of model 0 was 9414 ($-2\ell\ell$) and 9422 (Akaike's information criterion; AIC), respectively. Note that these values serve as a benchmark the fit of subsequent models can be compared to. Model 0 accounted for 37% of variance within persons (level 1).

Table 2. Parameter estimates.

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed effects						
α	75.27*	80.97*	75.20*	76.42*	77.98*	78.95*
γ	16.78*	17.05*	17.97*	17.76*	18.48*	17.93*
δ	14.24*	13.15*	14.07*	13.30*	13.06*	13.43*
Age $\rightarrow \alpha$			−2.29*	−1.75*	−1.63*	−1.57*
Age $\rightarrow \gamma$			1.11*	1.14*	1.16*	1.06*
Age $\rightarrow \delta$			0.55*	0.47*	0.46*	0.36*
Sex $\rightarrow \alpha$			4.61*	3.46*	4.04*	3.35*
Sex $\rightarrow \gamma$			−2.45*	−2.29*	−2.20*	−2.17*
Sex $\rightarrow \delta$			−0.75*	−0.71*	−0.84*	−0.77*
RecSongs $\rightarrow \alpha$				1.54*	1.57*	1.48*
RecSongs $\rightarrow \gamma$				0.18*	0.18*	0.17*
RecSongs $\rightarrow \delta$				0.39*	0.38*	0.37*
Taste Y $\rightarrow \alpha$					−2.76*	−2.45*
Taste Y $\rightarrow \gamma$					−1.05*	−1.02*
Taste Y $\rightarrow \delta$					0.45*	0.48*
ΔFreq $\rightarrow \alpha$						0.72*
ΔFreq $\rightarrow \gamma$						−0.31*
ΔFreq $\rightarrow \delta$						−0.46*
ΔPref $\rightarrow \alpha$						1.03*
ΔPref $\rightarrow \gamma$						−0.25*
ΔPref $\rightarrow \delta$						−0.36*
Random effects						
σ_e^2	711.14*	404.11*	386.05*	227.86*	223.75*	209.82*
σ_a^2		26.16*	25.16*	21.09*	20.84*	18.53*
σ_g^2		16.09*	15.25*	13.23*	12.20*	10.16*
σ_d^2		19.42*	18.73*	15.28*	15.09*	13.83*
Model fit						
$-2\ell\ell$	9414.2	9178.9	9115.9	8677.0	8668.8	8629.7
AIC	9422.2	9198.9	9147.9	8715.0	8712.8	8685.7
R_1^2	37%	64%	64%	64%	64%	64%
R_2^2	0%	0%	7%	31%	35%	42%

*$p < .05$.
Age is mean-centered; sex is coded as 0 = male, 1 = female; RecSongs = total number of recognized songs (mean-centred); taste Y = musical taste during a participant's youth, dummy-coded as 0 = "jazz/classical/rock/pop" and 1 = "German pop"; ΔFreq = listening frequency of bump over non-bump songs (mean-centred); ΔPref = preference of bump over non-bump songs (mean-centred); $-2\ell\ell$ = −2 times the log-likelihood (smaller is better); AIC = Akaike's information criterion (smaller is better); R_1^2 = explained variance at level 1 (within persons); R_2^2 = explained variance at level 2 (between persons).

In model 1, random effects for all three parameters were introduced. Moreover, covariances between these random effects were estimated. As can be seen from Table 2, the fixed effect estimate of maximum performance (α) increased somewhat after introducing random effects. All random effects variances for individual differences in α, γ, and δ were statistically significant, implying that there were reliable between-person differences in the "height" of the reminiscence bump, its location, and how pronounced it is.

In addition, two covariances (α with γ and δ) were significant. In a correlational metric, the association between individual differences in α and γ was $r = -0.29$, implying that those with a higher bump tended to show an earlier bump. The relation between α and δ was $r = 0.61$, showing that those with a higher bump also tended to show a more spread recognition performance across age. Compared to model 0, model 1 showed an improved fit. Moreover, model 1 explained 64% of variance within persons.

In the following model (model 2), age (mean-centred at 72.5 years) and sex were included as predictors. Age had a significant effect on all three parameters. The age effect was negative for maximum performance, implying that older participants had a significantly lower maximum performance. Also, the bump of older participants occurred later and their recognition performance was more spread across age. With respect to sex, women showed a bump about 2.5 years earlier than men. Compared to model 1, fit increased further. Because both predictors are on level 2, the amount of explained variance on level 1 remained unchanged, while age and sex accounted for 7% of variance between persons. Correlations among random effects were virtually the same as in model 1, but fixed estimates of α and δ changed somewhat due to scaling of the predictor variable sex. That is, for a male person of average age of the sample (i.e., 72.5 years), one would expect a maximum recognition performance of $\alpha = 75.2$ occurring at $\gamma = 17.97$ and a spreading of recognition performance of $\delta = 14.07$ (see Table 2).

For the next model (model 3), we added the number of recognized songs as a predictor variable. Table 2 displays that it had a positive effect on all three parameters, implying that those who recognized more songs had a higher bump, a later bump, and a more spread recognition performance. Including the number of recognized songs as a predictor greatly improved the model. The amount of explained variance on level 2 (between persons) increased and was estimated as 26%. The association between α and γ was no longer significant, while the correlation between α and δ was reduced.

Subsequently, in model 4, adolescent musical taste was included. As Table 2 shows, adolescent musical taste had a significant negative effect on γ, showing that those who mainly listened to “German pop” during their youth had an earlier bump (at the age of 17.43 years, on average) compared to those who preferred other musical genres during their youth. Compared to the previous model, fit increased slightly. Model 4 accounted for 64% and 31% of variance within and between persons, respectively.

In a final model (model 5), the frequency and preference differences between bump and non-bump songs were included as predictor variables. Table 2 indicates that both variables had statistically significant effects on α and γ. The effect on α was positive in both cases, implying that (1) the more often popular songs from one’s bump age are listened to and (2) the stronger the preference for bump songs is, the higher the maximum recognition performance, that is, the higher the bump. By contrast, the effect of frequency and preference differences was negative with respect to γ. Thus, (1) the more often participants listened to songs from their bump age and (2) the more they preferred songs from their bump age, the earlier the bump occurred. In comparison to model 4, fit improved further and the error variance decreased. Model 5 explained 64% and 39% of variance on level 1 and level 2, respectively. Compared to the previous model, the correlation between α and δ was reduced ($r = 0.35$).

In order to illustrate graphically the effect of the number of recognized songs and the amount of preference of bump songs over non-bump songs, Figure 3 depicts five curves. The continuous black curve represents the typical recognition performance curve across song-specific age. The light grey broken lines illustrate the effect of increasing or decreasing the number of recognized songs by one standard deviation. Analogously, the dark grey broken lines illustrate the effect of increasing or decreasing the preference of bump songs over non-bump songs by one

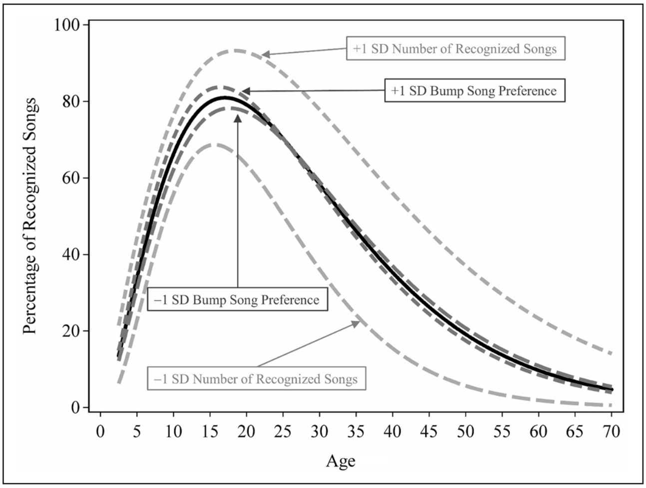

Figure 3. Effect of the number of recognized songs and of preference of bump songs over non-bump songs on the recognition performance curve. The black line represents the average recognition performance curve. The two light grey broken lines represent the recognition performance curves for participants one standard deviation above or below average in their number of recognized songs, respectively. The two dark grey broken lines represent the recognition performance curves for participants one standard deviation above or below average in their preference of bump over non-bump songs, respectively.

standard deviation. Figure 3 shows that the effect of the number of recognized songs is larger than the effect of bump song preference, because the former variable strongly affects the performance curve.

Discussion

The recognition of songs that have been popular in one's past is one manifestation of very long-term memory, that is, memory that stretches over years. What distinguishes long-term memory from other memory phenomena is that the former cannot be examined under the same experimental control. However, studies on very long-term memory usually allow for a more fine-graded performance measure, in this case, the percentage of recognized songs diagrammed across the song-specific age of participants, resulting in a *recognition performance curve*. Intriguingly, song recognition performance does not decline with the song-specific age of participants, but rather rises to a maximum during participants' adolescence – the so-called reminiscence bump (Schulkind et al., 1999; Zimprich & Wolf, 2016a). A possible reason for a bump in music recognition performance is that both personal and social identity develop during adolescence and both are linked to music preferences (Larson, 1995; North & Hargreaves, 1999; Schwartz & Fouts, 2003).

In the present study, our goal was to extend previous research on long-term memory for popular songs by examining whether the recognition performance curve (Figure 2) can be described by a non-linear power-exponential function. Our results show that the suggested function described recognition performance across the song-specific ages of participants very well.[6] What are the advantages of modeling a functional relation between recognition performance and song-specific age? One advantage is that a functional relation is more precise than a (verbal) description of recognition performance (Cavagnaro, Myung, & Pitt, 2013). Moreover, a functional relation is better suited for hypothesis testing (Luce, 1995). Finally, examining a functional relation comes closer to a model building and testing approach (Hedeker & Gibbons, 2006). We think that it is timely to integrate long-term memory research and advances in statistical modeling by examining the functional relation between recognition performance and song-specific age, much in the spirit of Borsboom (2006).[7]

The presence of a bump suggests that long-term retention is best for songs stemming from adolescence (Bartlett & Snelus, 1980; Schulkind et al., 1999). This could be due to the development of musical taste during adolescence (Hargreaves, North, & Tarrant, 2006; Holbrook & Schindler, 1989) and the time spent on music listening, which is maximal during adolescence (Bonneville-Roussy, Rentfrow, Xu, & Potter, 2013), and personal and social identity development (Larson, 1995; North & Hargreaves, 1999; Rentfrow, 2012). These different influences are difficult to disentangle because they develop together. Thus, the question of why a recognition performance bump occurs is difficult to answer – even more so because musical preferences and listening frequencies are assessed retrospectively.[8]

A second goal of the present study was to investigate individual differences in song recognition performance. There were reliable individual differences in all three parameters governing recognition curves. Thus, as is well known from other memory phenomena (e.g., Zimprich & Rast, 2009), individuals differ in terms of their performance. More specifically, individual differences were largest in the location of the bump and how pronounced the bump is.[9] Thus, what appears to be a unitary phenomenon, namely, a performance bump in music recognition, is relatively heterogeneous (Figure 3). Why are individual differences important? One reason is that individual differences *per se* offer new pathways for theoretical assumptions regarding long-term memory. As Underwood (1975, p. 129) has argued, “individual differences ought to be considered central in theory construction.” For example, instead of asking why there is a performance bump, one may ask why individuals differ in height, location, and shape of the bump (Zimprich & Wolf, 2016a, 2016b). In addition, the *amount* of individual differences in parameters and their correlations may be informative. In our data, the height of the bump differed less between persons than did its location and shape. Moreover, height and location were negatively correlated while height and shape were positively correlated, implying that those with an earlier bump had a lower bump and those with a higher bump showed a more “platykurtic” shape of the bump. Note that such a shift from group-based to individual-specific approaches parallels cognitive aging research, in which a number of studies provided new insights into cognitive aging by taking into account individual differences (e.g., Zimprich, 2002; Zimprich, Martin et al., 2008).

A third goal of the present study was to examine several predictors of very long-term memory for popular songs. The strongest predictor was the total number of correctly recognized songs. Those who recognized more songs showed a higher bump, a later bump, and a more “platykurtic” shape of the bump. In addition, demographic variables (age, sex), musical taste, and both the frequency and preference of listening to bump songs compared to non-bump songs emerged as predictors of at least one of the three parameters. Specifically, younger and female participants who recognized more songs in total and showed a higher listening frequency of and preference for bump songs had the highest bump. Younger and female participants who recognized fewer songs, described their adolescent musical taste as “German pop,” and showed a higher listening frequency of and preference for bump songs exhibited the earliest bump. Eventually, older participants who recognized more songs showed the most “platykurtic” shape of the bump (Zimprich & Wolf, 2016b). Predictors accounted for both *inter*individual differences and *intra*individual differences.

Although the age range of the present sample was small, an age effect emerged. Intriguingly, the age effect in γ was virtually 1, implying that with an increase of one year in age, the bump is shifted by one year. Obviously, this effect should not be extrapolated beyond the age range examined here, because it may lead to meaningless predictions. With respect to sex, women exhibited an earlier bump. One explanation for this finding is that girls enter puberty earlier and, hence, may be interested in music earlier. In our study, the total number of recognized songs was the most powerful predictor of the recognition performance curve. Although it accounted for about 24% of individual differences, this also implies that participants who recognized the same number of songs exhibit different recognition curves (Schellenberg & Habashi, 2015). In line with Holbrook and Schindler (1989), we found that the preferred music styles during people's youth affected recognition. Indeed, adolescent musical taste affected the location of the bump – those with a preference for "German pop" had an earlier bump, probably because many of the stimuli songs belonged to this genre. As in Janssen and colleagues (2007), the "rehearsal" (frequency of listening) of songs was one explanation for the high recognition rate of bump songs. Similarly, preference affected both height and location of the bump. However, because participants received feedback on whether they correctly recognized a song before being asked for listening frequency and preference, a confounding may have occurred.

There are other predictors that may explain individual differences in the song recognition performance curve. One may distinguish between predictors based on maximum performance and predictors based on typical performance. If popular song recognition were similar to more typical memory phenomena, one would expect that predictors based on maximum performance, e.g., cognitive variables like processing speed (e.g., Zimprich, Rast, et al., 2008), would account for substantial amounts of variance. However, very long-term memory for music appears to be also affected by predictors based typical performance. For example, openness for experience tends to decrease across the lifespan (Zimprich, Allemand, & Lachman, 2012), which may result in a tendency to re-listen to preferred music instead of "discovering" new music. In line with this assumption, Rawlings and Ciancarelli (1997) have shown that more open individuals liked a wider range of music types.

To conclude, the present study shows the benefits of a (non-linear) functional relation between music recognition performance and the song-specific age of participants. Moreover, taking into account individual differences has the potential to increase further our understanding of very long-term memory. The approach suggested here can, *mutatis mutandis*, be transferred to other types of stimuli but also to the direct examination of the recall of autobiographical events (Wolf & Zimprich, 2015).

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Here, we confine ourselves to songs. But similar findings, with a grain of salt, have been reported for other types of stimuli (e.g., Janssen et al., 2006).
2. A similar account based on identity-formation has been proposed to explain the reminiscence bump in autobiographical memory (e.g., Rathbone, Moulin, & Conway, 2008).
3. In Zimprich and Wolf (2016a), the goal was to model individual differences in the distribution of correctly recognized songs. In the present paper, the goal is to establish individual differences in a functional relation between age (when a song was popular) and performance, that is, correct recognition of songs. Consequently, different analytical approaches were used.

4. To estimate the amount of between-person and within-person variance, we, additionally, estimated an intercept-only model. In this model, the intercept variance estimate was 253.18, while the error variance estimate was 882.02. Together, this leads to an intraclass correlation coefficient of $r_{icc} = \frac{\text{between-person variance}}{\text{total variance}} = \frac{253.18}{253.18+882.02} = 22\%$, implying that 22% of the total variance was between persons and, thus, 78% of the total variance within persons.
5. To achieve this result, one calculates $\gamma+\delta=16.78+14.24=31.02$ years.
6. Note that, naturally, there are alternatives to the combined power-exponential function used in the present research. The Hailwood–Horrobin or inverse second-order polynomial function (Nelder, 1966) might be a good candidate. In a direct comparison, however, a non-linear model based on the Hailwood–Horrobin function achieved a worse fit ($-2\ell\ell = 9488.0$; AIC = 9496; $R^2 = 32\%$) compared to model 0 from Table 2. Apart from the functional form, the distribution of errors may differ from normal. In our data, because performance is expressed as percentages, there is an upper and lower boundary of possible values, which, strictly speaking, makes a normal distribution of errors implausible – at least for values close to the boundaries. An alternative to the normal distribution might be the beta distribution (Zimprich, 2010). A direct comparison, though, showed that a model based on the beta distribution evinced an inferior fit ($-2\ell\ell = 10419$; AIC = 10427; $R^2 = 28\%$) compared to model 0 from Table 2 – probably because predicted values are not close to the boundaries (see Figure 2).
7. To begin with, future studies could – much like research on the functional form of the learning or the forgetting curve – focus on the question of whether a combined power-exponential function also describes other data sets adequately or whether there are better alternatives. That is, it is not so much the specific parameter estimates that may be the target of replication studies – these are expected to differ in dependence on stimulus material, experimental design, etc. Rather, it is the functional form itself that should be gauged.
8. An anonymous reviewer wondered about the high rate of recognition of "bump songs," which most likely was due to the selection of stimuli songs. Because the present study, however, focused on individual differences in the recognition curve, we intended to choose popular songs to increase the likelihood that most participants had heard most songs before. And while this may have elevated the average recognition curve as a whole, at the same time it increased individual differences.
9. To see this, note that the coefficient of variation (CV) or relative standard deviation (RSD) is defined as the ratio of the standard deviation σ to the mean μ, that is, $\mathrm{CV} = \mathrm{RSD} = \sigma/\mu$. Inserting the according values of model 1 from Table 2, one achieves $CV\alpha = 0.068$, $CV\gamma = 0.239$, and $CV\delta = 0.309$, showing that the RSD is much larger for γ and δ.

References

Bartlett, J. C., & Snelus, P. (1980). Lifespan memory for popular songs. *American Journal of Psychology*, *93*, 551–560.

Bonneville-Roussy, A., Rentfrow, P. J., Xu, K. M., & Potter, J. (2013). Music through the ages: Trends in musical attitudes and preferences from adolescence through middle adulthood. *Journal of Personality and Social Psychology*, *105*, 703–717.

Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika*, *71*, 425–440.

Cavagnaro, D. R., Myung, J. I., & Pitt, M. A. (2013). Mathematical modeling. In T. Little (Ed.), *Oxford handbook of quantitative methods, Vol. 1* (pp. 438–553.). Oxford: Oxford University Press.

Conway, M. A. (2005). Memory and the self. *Journal of Memory and Language*, *53*, 594–628.

Davidian, M., & Giltinan, D. M. (1995). *Nonlinear models for repeated measurement data*. New York: Chapman & Hall.

Erikson, E. H. (1950). *Childhood and society* (1st ed.). New York: Norton.

Hargreaves, D. J., North, A. C., & Tarrant, M. (2006). Musical preference and taste in childhood and adolescence. In G. E. McPherson (Ed.), *The child as musician* (pp. 135–154). Oxford: Oxford University Press.

Hébert, S., & Peretz, I. (1997). Recognition of music in long-term memory: Are melodic and temporal patterns equal partners? *Memory and Cognition*, *25*, 518–533.

Hedeker, D., & Gibbons, R. D. (2006). *Longitudinal data analysis*. New York: Wiley.
Holbrook, M. B., & Schindler, R. M. (1989). Some exploratory findings on the development of musical tastes. *Journal of Consumer Research*, *16*, 119–124.
Janssen, S. M. J., Chessa, A. G., & Murre, J. M. J. (2007). Temporal distribution of favourite books, movies, and records: Differential encoding and re-sampling. *Memory*, *15*, 755–767.
Janssen, S. M. J., Murre, J. M. J., & Meeter, M. (2008). Reminiscence bump in memory for public events. *European Journal of Cognitive Psychology*, *20*, 738–764.
Kurtz, T., & Zimprich, D. (2013). Individual differences in criterion-based dropout learning in old age – The role of processing speed and verbal knowledge. *European Journal of Ageing*, *11*, 183–193.
Larsen, S. F. (1996). Memorable books: Recall of reading and its personal context. In R. Kreuz & M. S. MacNealy (Eds.), *Empirical approaches to literature and aesthetics* (pp. 585–599). Norwood, NJ: Ablex.
Larson, R. (1995). Secrets in the bedroom: Adolescents' private use of media. *Journal of Youth and Adolescence*, *24*, 535–550.
Lonsdale, A. J., & North, A. C. (2011). Why do we listen to music? A uses and gratifications analysis. *British Journal of Psychology*, *102*, 108–134.
Luce, R. D. (1995). Four tensions concerning mathematical modeling in psychology. *Annual Review of Psychology*, *46*, 1–26.
Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, *55*, 107–122.
Molenberghs, G., & Verbeke, G. (2005). *Models for discrete longitudinal data*. New York: Springer.
Nelder, J. A. (1966). Inverse polynomials, a useful group of multi-factor response functions. *Biometrics*, *22*, 128–141.
Neisser, U. (1978). Memory: What are the important questions? In M. M. Gruneberg, P. E. Morris & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 3–24). London: Academic Press.
North, A. C., & Hargreaves, D. J. (1999). Music and adolescent identity. *Music Education Research*, *1*, 75–92.
Platz, F., Kopiez, R., Hasselhorn, J., & Wolf, A. (2015). The impact of song-specific age and affective qualities of popular songs on music-evoked autobiographical memories (MEAMs). *Musicae Scientiae*, *19*, 327–349.
Rathbone, C. J., Moulin, C. J. A., & Conway, M. A. (2008). Self-centred memories: The reminiscence bump and the self. *Memory & Cognition*, *36*, 1403–1414.
Rawlings, D., & Ciancarelli, V. (1997). Music preference and the five-factor model of the NEO personality inventory. *Psychology of Music*, *25*, 120–132.
Rentfrow, P. J. (2012). The role of music in everyday life: Current directions in the social psychology of music. *Social and Personality Psychology Compass*, *6*, 402–416.
Rubin, D. C. (1977). Very long-term memory for prose and verse. *Journal of Verbal Learning and Verbal Behavior*, *16*, 611–621.
Rubin, D. C., Rahhal, T. A., & Poon, L. W. (1998). Things learned in early adulthood are remembered best. *Memory & Cognition*, *26*, 3–19.
Rubin, D. C., & Schulkind, M. D. (1997). The distribution of autobiographical memories across the lifespan. *Memory & Cognition*, *25*, 859–866.
SAS Institute Inc. (2008). *SAS/STAT 9.2 user's guide*. Cary, NC: SAS Institute Inc.
Schellenberg, E. G., & Habashi, P. (2015). Remembering the melody and timbre, forgetting the key and tempo. *Memory & Cognition*, *43*, 1021–1031.
Schindler, R. M. & Holbrook, M. B. (2003). Nostalgia for early experience as a determinant of consumer preferences. *Psychology & Marketing, 20*, 275–302.
Schulkind, M. D., Hennis, L. K., & Rubin, D. C. (1999). Music, emotion, and autobiographical memory: They're playing your song. *Memory & Cognition*, *27*, 948–955.
Schwartz, K. D., & Fouts, G. T. (2003). Music preferences, personality style, and developmental issues of adolescents. *Journal of Youth and Adolescence*, *32* (3), 205–213.
Snijders, T. A. B., & Bosker, R. J. (1994). Modeled variance in two-level models. *Sociological Methods and Research*, *22*, 342–363.
Underwood, B. J. (1975). Individual differences as a crucible in theory construction. *American Psychologist*, *30*, 128–134.

Wolf, T., & Zimprich, D. (2015). How can individual differences in autobiographical memory distributions of older adults be explained? *Memory*, *24*(9), 1287–1299. doi:10.1080/09658211.2015.1102291
Zimprich, D. (2002). Cross-sectionally and longitudinally balanced effects of processing speed on intellectual abilities. *Experimental Aging Research*, *28*, 231–251.
Zimprich, D. (2010). Modeling change in skewed variables using mixed beta regression models. *Research in Human Development*, *7*, 9–26.
Zimprich, D., Allemand, M., & Lachman, M. E. (2012). Factorial structure and age-related psychometrics of the MIDUS personality adjective items across the lifespan. *Psychological Assessment*, *24*, 173–186.
Zimprich, D., & Kurtz, T. (2013). Individual differences and predictors of forgetting in old age: The role of processing speed and working memory. *Aging, Neuropsychology, & Cognition*, *20*, 195–219.
Zimprich, D., Martin, M., Kliegel, M., Dellenbach, M., Rast, P., & Zeintl, M. (2008). Cognitive abilities in old age: Results from the Zurich Longitudinal Study on Cognitive Aging. *Swiss Journal of Psychology*, *67*, 177–195.
Zimprich, D., & Rast, P. (2009). Verbal learning changes in older adults across 18 months. *Aging, Neuropsychology, and Cognition*, *16*, 461–484.
Zimprich, D., Rast, P., & Martin, M. (2008). Individual differences in verbal learning in old age. In S. M. Hofer & D. F. Alwin (Eds.), *The handbook of cognitive aging: Interdisciplinary perspectives* (pp. 224–243). Thousand Oaks, CA: Sage Publications.
Zimprich, D., & Wolf, T. (2016a). The distribution of memories for popular songs in old age: An individual differences approach. *Psychology of Music*, *44*, 640–657.
Zimprich, D., & Wolf, T. (2016b). Modeling individual differences in autobiographical memory distributions using mixed logitnormal regression. *Applied Cognitive Psychology*, *30*, 360–374.

Appendix. Stimulus songs.

Year	Artist – song
1945	Kitty Kallen – “It’s been a long, long time”
1946	Willi Kollo – “Lieber Leierkastenmann”
1947	Maria Andergast – “Mariandl”
1948	Theo Lingen – “Der Theodor im Fußballtor”
1949	Lonny Kellner – “Im Hafen von Adano”
1950	Rene Carol – “Buonanotte angelo mio”
1951	Les Paul and Mary Ford – “How high the moon”
1952	Zarah Leander – “Wunderbar”
1953	Maria von Schmedes – “Ich möcht gern dein Herz klopfen hörn”
1954	Illo Schieder und Max Greger – “Sieben einsame Tage”
1955	The McGuire Sisters – “Sincerely”
1956	Angèle Durand – “So ist Paris”
1957	Caterina Valente – “Dich werde ich nie vergessen”
1958	Edmundo Ros – “Melodie d’amour”
1959	Hazy Osterwald – “Kriminaltango”
1960	Ted Herold – “Moonlight”
1961	Connie Francis – “Schöner fremder Mann”
1962	Conny Froebess – “Zwei kleine Italiener”
1963	Peter Alexander – “Wenn erst der Abend kommt”
1964	Peter Lauch – “Das kommt vom Rudern”
1965	Cliff Richard – “Das ist die Frage aller Fragen”
1966	Chris Andrews – “To whom it concerns”
1967	Manfred Mann – “Haha said the clown”
1968	The Tremoloes – “My little lady”

Appendix. (Continued)

Year	Artist – song
1969	Elvis Presley – "In the ghetto"
1970	The Kinks – "Lola"
1971	Roy Black – "Für Dich allein (Du kannst nicht alles haben)"
1972	The Sweet – "Poppa Joe"
1973	The Sweet – "Blockbuster"
1974	Nick MacKenzie – "Juanita"
1975	Howard Carpendale – "Deine Spuren im Sand"
1976	Sailor – "A glass of champagne"
1977	Smokie – "Lay back in the arms of someone"
1978	John Paul Young – "Love is in the air"
1979	Nick Straker Band – "A walk in the park"
1980	Roland Kaiser – "Santa Maria"
1981	Kim Wilde – "Kids in America"
1982	Hubert Kah – "Rosemarie"
1983	New Order – "Blue Monday"
1984	Limahl – "The neverending story"
1985	Harold Faltermeyer – "Axel F"
1986	Bruce & Bongo – "Geil"
1987	Jürgen von der Lippe – "Guten Morgen, liebe Sorgen"
1988	Guillermo Marchena – "My love is a tango"
1989	The Cure – "Lullaby"
1990	Twenty 4 Seven – "I can't stand it"
1991	Crystal Waters – "Gypsy woman"
1992	Jon Secada – "Just another day"
1993	DJ BoBo – "Somebody dance with me"
1994	Prince Ital Joe Feat, Marky Mark – "United"
1995	Take That – "Back for good"